VECTOR ANALYSIS

WITH AN INTRODUCTION TO

TENSOR ANALYSIS

BY

A. P. WILLS

LATE PROFESSOR OF MATHEMATICAL PHYSICS
COLUMBIA UNIVERSITY

First Printing ....October 1931
Second Printing ....October 1938
Third Printing ....June 1940
Fourth Printing ....June 1947
Fifth Printing ....May 1949
Sixth Printing ....April 1950

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

IN writing the present book the author has endeavored to present the subject matter in a manner which might appeal particularly to students who are specially interested in the study of Physics.

The treatment of vectors of 3-dimensional Euclidean space is roughly equivalent to that which can be found in several other English texts. The theory of vectors and tensors associated with non-Euclidean metrical manifolds is given by methods which, so far as the author is aware, are not to be found elsewhere in English. These methods, whose characteristic features are due to the German mathematician Gerhard Hessenberg, simplify materially the treatment of the theory of tensors.

Geometrical and physical applications are given in sufficient number to exemplify the methods of the text, and are not in general such as to necessitate the assumption of special mathematical and physical knowledge on the part of the reader.

To his colleagues, Dr. T. H. Gronwald and Dr. S. Serghiesco, the author wishes to express his appreciation—to the former for looking over the original manuscript and for various valuable suggestions, and to the latter for his kind assistance in the proof reading. Acknowledgments are also gratefully made by the author to the publishers and printers for their patient and helpful collaboration in handling the technical details connected with the preparation of the book.

A. P. WILLS.

August, 1931.

CONTENTS

Chapter V

✓ SCALAR AND VECTOR POTENTIAL FUNCTIONS

Chapter VI

LINEAR VECTOR FUNCTIONS AND DYADICS

ARTICLE PAGE

Chapter IX

NON-EUCLIDEAN MANIFOLDS

Chapter X

TENSOR THEORY

HISTORICAL INTRODUCTION

THE origin of the branch of mathematics now known as Vector Analysis can be traced to early attempts to find geometric representations of imaginary algebraic quantities.

Long before these attempts were made, mathematicians had recognized the convenience of representing positive and negative quantities by the distances laid off in opposite directions on a straight line from a fiducial point called the origin; this straight line would now be called the Axis of Reals.

In his *Algebra*, published in 1673, John Wallis (1616–1703), Savilian professor (1649–1703) at Oxford, showed how to represent geometrically impossible (complex) roots of a quadratic equation; in his representation, a distance from the origin along the axis of reals represented the real part of the root, the distance being measured in the positive or negative direction of the axis according as this part of the root was positive or negative; from the point on the real axis thus determined, a line was drawn perpendicular to the axis of reals whose length represented the number which multiplied by $\sqrt{-1}$ gave the imaginary part of the root, the line being drawn in one direction or the opposite according as the number was positive or negative.

From the construction of Wallis to the introduction of an Axis of Imaginaries perpendicular to the real axis with $\sqrt{-1}$ as an associated unit does not appear to us now as a very great step, yet it was apparently not definitely taken until more than a century after Wallis published his *Algebra.*

In 1798 the Norwegian surveyor Caspar Wessel (1745–1818) wrote a paper which appeared in the *Proceedings of the Royal Society of Denmark* in 1799, entitled *On the Analytic Representation of Direction; an Attempt.* In this paper an axis of imaginaries with $\sqrt{-1}$ as an associated unit is definitely introduced.

After defining the addition of right lines by a method completely analogous to that for the addition of forces, Wessel proceeds to define the product of two right lines in a plane by means of the following rules[1]:

[1] From a translation by Professor Martin A. Nordgaard in Professor David Eugene Smith's *Source Book of Mathematics*, published in 1929 by McGraw-Hill Book Co., Inc.

"Firstly, the factors shall have such a direction that they both can be placed in the same plane with the positive unit.

"Secondly, as regards length, the product shall be to one factor as the other factor is to the unit.

"Finally, if we give the positive unit, the factors, and the product, a common origin, the product shall, as regards its direction, lie in the plane of the unit and the factors, and diverge from the one factor as many degrees, and on the same side, as the other factor diverges from the unit, so that the direction angle of the product, or its divergence from the positive unit, becomes equal to the sum of the direction angles of the factors."

He then proceeds as follows:

"Let $+1$ designate the positive rectilinear unit and ϵ a certain other unit perpendicular to the positive unit and having the same origin; then the direction angle of $+1$ will be equal to $0°$, that of -1 to $180°$, that of $+\epsilon$ to $90°$, and that of $-\epsilon$ to $-90°$ or $270°$. By the rule that the direction angle of the product shall equal the sum of the angles of the factors, we have:

$$(+1)(+1) = +1;\quad (+1)(-1) = -1;\quad (-1)(-1) = +1;$$
$$(+1)(+\epsilon) = +\epsilon;\quad (+1)(-\epsilon) = -\epsilon;\quad (-1)(+\epsilon) = -\epsilon;$$
$$(-1)(-\epsilon) = +\epsilon;\quad (+\epsilon)(+\epsilon) = -1;\quad (+\epsilon)(-\epsilon) = +1;$$
$$(-\epsilon)(-\epsilon) = -1.$$

"From this it is seen that $\epsilon = \sqrt{-1}$; and the divergence of the product is determined such that not any of the common rules of operation are contravened."

On Wessel's scheme of representation the terminal point P of a right line (line-vector) OP drawn from the origin O in the plane of the units $+1$ and $\sqrt{-1}$ represents a complex number, say $a + b\sqrt{-1}$; and, similarly, the terminal point Q of a line-vector OQ represents another complex number, say $c + d\sqrt{-1}$. In these representations, a, b and c, d, respectively, may be considered as plane rectangular Cartesian co-ordinates with respect to Cartesian axes which coincide with those associated with the unit $+1$ and the unit $\sqrt{-1}$.

The complex numbers $a + b\sqrt{-1}$ and $c + d\sqrt{-1}$ can be said to specify the line-vectors OP and OQ, respectively. The product of the two line-vectors is a new line-vector, say OS, also in the plane of the units, specified by a complex number which is the product of those specifying OP and OQ and which is obtained as follows:

$$(a + b\sqrt{-1})(c + d\sqrt{-1}) = ac - bd + (ad + bc)\sqrt{-1}.$$

The relationship of the product to the factors is better stated through the trigonometric expressions for the complex numbers which specify them, or better still through the corresponding exponential expressions. We can write:

$$a + b\sqrt{-1} = r(\cos\theta + \sqrt{-1}\sin\theta) = re^{\sqrt{-1}\theta},$$
$$c + d\sqrt{-1} = s(\cos\phi + \sqrt{-1}\sin\phi) = se^{\sqrt{-1}\phi}$$

where r denotes the length of the line-vector OP and θ the angle which it makes with the positive direction of the axis of reals, and s and ϕ have similar meanings for the line-vector OQ. The third equation back can then be written in the form:

$$(re^{\sqrt{-1}\theta})(se^{\sqrt{-1}\phi}) = rse^{\sqrt{-1}(\theta+\phi)}.$$

Here, the expression on the right specifies the line-vector product of the line-vectors specified by the factors on the left. The line-vector product lies in the plane of the factors, has a length equal to the product of the lengths of the factors, and makes an angle with the positive direction of the axis of reals which is equal to the sum of the angles made by the factors with the same direction.

It follows that any one of the equivalent expressions

$$a + b\sqrt{-1}, \quad r(\cos\theta + \sqrt{-1}\sin\theta), \quad re^{\sqrt{-1}\theta}$$

can be considered as a plane operator which has the property of rotating a line-vector drawn from the origin in the plane of the real and imaginary units through the angle θ in the direction reckoned as positive, and stretching it in the ratio $r:1$.

These results have found numerous physical applications, particularly in the theory of alternating currents.

The method of Wessel, before the publication of the French translation of his paper in 1897, was commonly attributed to other writers, in particular to J. R. Argand, whose name is often associated with Wessel's construction.

Wessel sought to extend his method to space of three dimensions, and had he been successful would very likely have been led to the invention of quaternions.

This unsolved problem was taken up by the French mathematician Servois, whose results were published in 1813 in Gergonne's *Annales*. He sought to find by analogy from the complex expression $a + b\sqrt{-1}$ for a line-vector in a plane, a corresponding expression for a line-vector in space of three dimensions, and suggested for a unit line-vector the form:

$$p \cos \alpha + q \cos \beta + r \sin \gamma,$$

where α, β, γ are its direction angles with respect to three axes, and where p, q, r are non-real quantities which he thought might be reducible to the general form $A + B\sqrt{-1}$. This, however, he was not able to prove, but it turned out later that the p, q, r of Servois are equivalent to three of the fundamental units in Hamilton's theory of Quaternions.

The invention of Quaternions was announced by Sir William Rowan Hamilton (1805–1865), professor of astronomy in the University of Dublin and Royal Astronomer of Ireland, at a meeting of the Royal Irish Academy in November, 1843.

The specification of a quaternion requires in general four real parameters, the unit $+1$ of real numbers, and three other units, **i**, **j**, **k**, whose properties will be given presently. That four real parameters are required in general for the specification, follows from the property which a quaternion must possess when considered as an operator, analogous to the plane operator $a + b\sqrt{-1}$, of rotating a line-vector through a given angle about an axis through its initial point and stretching it in a given ratio; for two parameters are required to specify the axis, one to specify the angle of rotation, and one to specify the stretch-ratio, making four parameters in all.

The fundamental units **i**, **j**, **k** on Hamilton's quaternionic scheme are subject to the following rules of multiplication:

$$\begin{aligned} \mathbf{jk} &= -\mathbf{kj} = \mathbf{i}; \\ \mathbf{ki} &= -\mathbf{ik} = \mathbf{j}; \\ \mathbf{ij} &= -\mathbf{ji} = \mathbf{k}; \\ \mathbf{i}^2 &= \mathbf{j}^2 = \mathbf{k}^2 = -1. \end{aligned}$$

It will be noticed that for these units the commutative law of multiplication has been abandoned.

A quaternion **q** can be expressed in the form:

$$\mathbf{q} = \omega + a\mathbf{i} + b\mathbf{j} + c\mathbf{k}.$$

It consists of two parts: a Scalar ω and a Vector

$$\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k},$$

where a, b, c are rectangular Cartesian co-ordinates of a point P and the units **i**, **j**, **k** represent unit vectors in the positive directions of the corresponding axes. This vector specifies a line-vector from the origin O to the point P.

The criterion of equality of two quaternions is that their scalar parts shall be equal and that the coefficients of their **i**, **j**, **k**-units shall be respectively equal.

Two quaternions are added by adding their scalar parts and adding the coefficients of each of the **i**, **j**, **k**-units to form new coefficients for those units. The sum of two quaternions is therefore itself a quaternion.

All the familiar algebraic rules of multiplication are supposed valid in operating with quaternions, except that in forming products of the units **i**, **j**, **k** the rules given above, which do not include the commutative law, must be used.

When these rules are borne in mind, the product of two quaternions can easily be shown to be a quaternion. The product of a quaternion $\omega + a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ and a vector $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is of special interest. It is expressed as follows:

$$
\begin{aligned}
(\omega + a\mathbf{i} + b\mathbf{j} + c\mathbf{k})(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = & -(ax + by + cz) \\
& +(\omega x + bz - cy)\,\mathbf{i} \\
& +(\omega y + cx - az)\,\mathbf{j} \\
& +(\omega z + ay - bx)\,\mathbf{k}.
\end{aligned}
$$

By a proper choice of the four parameters ω, a, b, c, this quaternion can be reduced to any assigned vector $x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}$; the criterion of equality of two quaternions requires that the coefficients of each of their four units, $+1$, **i**, **j**, **k**, shall be respectively equal; hence, to effect the reduction we must have:

$$
\begin{aligned}
ax + by + cz &= 0, \\
\omega x + bz - cy &= x', \\
\omega y + cx - az &= y', \\
\omega z + ay - bx &= z'.
\end{aligned}
$$

These four equations suffice to determine the four parameters ω, a, b, c, in terms of x, y, z and x', y', z'.

This result is of special importance since it supplies the operator, viz: the quaternion $\omega + a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, which will rotate a line-vector through a given angle about a given axis in space through its initial point and stretch it in a given ratio.

By forming on Hamilton's scheme the product of two vectors, a quaternion results, thus:

$$
\begin{aligned}
(a\mathbf{i} + b\mathbf{j} + c\mathbf{k})(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = & -(ax + by + cz) \\
& +(bz - cy)\,\mathbf{i} \\
& +(cx - az)\,\mathbf{j} \\
& +(ay - bx)\,\mathbf{k};
\end{aligned}
$$

the negative of the scalar part of this quaternion would now be called the scalar product, and the vector part the vector product of the two vectors.

A scalar function $u(x, y, z)$, where x, y, z are the co-ordinates of a point $P(x, y, z)$ in space, is called a scalar point function. Hamilton introduced the corresponding idea of a vector point function, say $\mathbf{v}(x, y, z)$, having a value at each point $P(x, y, z)$ of space, depending on the position of the point.

In his development of the calculus of quaternions Hamilton was led to the invention of the important differential operator:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}.$$

If u denote a continuous scalar point function, then:

$$\begin{aligned}\nabla u &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) u \\ &= \mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z};\end{aligned}$$

and ∇ acting upon a continuous scalar point function u therefore produces a vector point function, now known as the gradient of u; it represents in magnitude and direction the greatest space rate of increase of u.

If $\mathbf{v}(= v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k})$ denote a continuous vector point function, then:

$$\begin{aligned}\nabla\mathbf{v} &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) \\ &= -\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}\right) \\ &\quad + \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} \\ &\quad + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} \\ &\quad + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k};\end{aligned}$$

and the result of operating with ∇ upon a continuous vector point function $\mathbf{v}$ is to produce a quaternion; the scalar part of this quaternion is the negative of what is now called the divergence of $\mathbf{v}$, and the vector part is what is now called the curl of $\mathbf{v}$.

In spite of the many beautiful and suggestive results brought forward by Hamilton in his development of the algebra and calculus of quaternions, a few of which have been noticed above, the quaternionic scheme failed to meet in an entirely satisfactory way the requirements of mathematical physicists, notwithstanding able advocates.

Hamilton, in his *Lectures on Quaternions* (1853), developed the subject from a geometric point of view, but there is little doubt that his invention of quaternions was the result of algebraic rather than geometric reasoning, in following which he was actually engaged in the development of a quadruple algebra.

Vector and tensor algebra can be considered as a branch of multiple algebra, their common origin being found in the early attempts at geometrical representation of complex numbers, in which are involved the fundamental notions of double algebra, as the method of Wessel, outlined above, clearly shows. Long before the invention of quaternions, August Ferdinand Möbius (1790–1868) published (1827) his remarkable treatise on *Der Barycentrische Calcul* in which the initial steps were taken in the development of a quadruple algebra, but he was unable to discover appropriate rules to govern the formation of products of multiple quantities, such as Hamilton found later for his quaternions and Grassmann for his *Extensive Grössen.*

In the year 1844, the next following that in which Hamilton announced his discovery of quaternions, Hermann Grassmann (1809–1877), professor of mathematics in the gymnasium at Stettin, Germany, published the first edition of his celebrated treatise on space analysis, entitled *Die Lineale Ausdehnungslehre, ein neuer Zweig der Mathematik.* In this edition he developed a non-metrical geometry of points, applicable to space of any number of dimensions, from a somewhat abstract philosophical point of view which, apparently, was distasteful to the mathematicians of his time, for in spite of its remarkable originality and suggestiveness the book attracted little attention for many years following its publication. On this account, probably, the book was revised and rewritten in more readable form and a second edition, entitled *Die Ausdehnungslehre, vollständig und in strenger Form bearbeitet,* was published in 1862. Our discussion here must be confined to mention of that part of Grassmann's great work which is specifically concerned with the algebra of vectors and tensors.

At the beginning of the second edition of the *Ausdehnungslehre,* Grassmann introduced the notion of a sort of hyper-number which he called an *Extensive Grösse.* An example of such a hyper-number, called a primary hyper-number, is furnished by the polynomial

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdot \cdot \cdot + \alpha_n\mathbf{e}_n = \sum_r^n \alpha_r\mathbf{e}_r,$$

where the α-coefficients denote ordinary numbers and the $\mathbf{e}$'s certain

primary units upon which, as will be explained below, various conditions may be imposed.

Two such primary hyper-numbers are added as follows:

$$\sum_r^n \alpha_r \mathbf{e}_r + \sum_r^n \beta_r \mathbf{e}_r = \sum_r^n (\alpha_r + \beta_r)\, \mathbf{e}_r.$$

Multiplication and division of hyper-numbers by ordinary numbers are subject to the laws of ordinary algebra.

The general product of two primary hyper-numbers is formed as follows:

$$\sum_r^n \alpha_r \mathbf{e}_r \sum_r^n \beta_s \mathbf{e}_s = \sum_{r,s}^n \alpha_r \beta_s \mathbf{e}_r \mathbf{e}_s,$$

where the products $\mathbf{e}_r\mathbf{e}_s$ of two primary units are called units of the second order, the number of such units being equal to n^2. Upon these units special conditions may be imposed in a variety of ways, and according to the nature of such conditions special products may be derived, among which here are two of special importance in vector and tensor algebra, viz: the Inner Product and the Outer or Combinatory Product.

In the case of an inner product the units of the second order are subject to the following conditions:

$$\mathbf{e}_r|\mathbf{e}_r = 1, \quad \mathbf{e}_r|\mathbf{e}_s = 0, \quad (s \neq r),$$

where the vertical line between two units indicates that their inner product is to be understood.

In the case of an outer product the conditions to which the units of the second order are subject are as follows:

$$[\mathbf{e}_r\mathbf{e}_r] = 0, \quad [\mathbf{e}_r\mathbf{e}_s] = -[\mathbf{e}_s\mathbf{e}_r],$$

where the square brackets enclosing two primary units indicate that an outer product is to be understood.

The significance of such products (and of others of higher order formed in similar ways) with relation to the affine geometry of a space of n-dimensions was defined and explained in great detail by Grassmann. For the purposes of the present review it will suffice to consider the significance of the inner and outer products of two hyper-numbers representing vectors in space of 3-dimensions.

Consider the two primary hyper-numbers:

$$\alpha = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3,$$
$$\beta = \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \beta_3\mathbf{e}_3,$$

where $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are primary units, represented geometrically by directed line-segments of unit length drawn from a common origin so as to determine a right-handed orthogonal system of axes; $\alpha_1\mathbf{e}_1$, $\alpha_2\mathbf{e}_2$, $\alpha_3\mathbf{e}_3$ are multiples of the primary units, represented geometrically by orthogonal projections upon these axes of a directed line-segment in space representing $\boldsymbol{\alpha}$, drawn from the origin, and $\beta_1\mathbf{e}_1$, $\beta_2\mathbf{e}_2$, $\beta_3\mathbf{e}_3$ have a corresponding significance for a second directed line-segment representing $\boldsymbol{\beta}$; the term *Strecke*, for which the term line-vector will be substituted, was used by Grassmann to designate a directed line-segment.

With the aid of the inner product rules for primary units stated above the inner product of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be expressed as follows:

$$\boldsymbol{\alpha}|\boldsymbol{\beta} = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3.$$

Evidently:

$$\boldsymbol{\beta}|\boldsymbol{\alpha} = \boldsymbol{\alpha}|\boldsymbol{\beta}.$$

The numerical value or magnitude of a hyper-number is defined as the positive square root of the inner product of the hyper-number by itself. Hence, if α and β denote the magnitudes of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ respectively:

$$\alpha = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}, \quad \beta = \sqrt{\beta_1^2 + \beta_2^2 + \beta_3^2}.$$

The magnitudes of the hyper-numbers $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are therefore numerically equal to the lengths of the line-vectors which represent them geometrically. If θ denote the angle between these line-vectors, then:

$$\boldsymbol{\alpha}|\boldsymbol{\beta} = \alpha\beta\left(\frac{\alpha_1}{\alpha}\frac{\beta_1}{\beta} + \frac{\alpha_2}{\alpha}\frac{\beta_2}{\beta} + \frac{\alpha_3}{\alpha}\frac{\beta_3}{\beta}\right) = \alpha\beta\cos\theta.$$

With the aid of the outer product rules for units of the second order the outer product $\mathbf{P}$ of the hyper-numbers $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be expressed as follows:

$$\begin{aligned}\mathbf{P} = [\boldsymbol{\alpha\beta}] = \; & (\alpha_2\beta_3 - \alpha_3\beta_2)\,[\mathbf{e}_2\mathbf{e}_3] \\ & +(\alpha_3\beta_1 - \alpha_1\beta_3)\,[\mathbf{e}_3\mathbf{e}_1] \\ & +(\alpha_1\beta_2 - \alpha_2\beta_1)\,[\mathbf{e}_1\mathbf{e}_2].\end{aligned}$$

This product is a hyper-number of the second order and is expressed in terms of independent units of the second order. Its numerical value (magnitude) P is defined as $\sqrt{\mathbf{P}|\mathbf{P}}$; the definition of the inner product of two hyper-numbers of the second order is such as to give:

$$P = \sqrt{\mathbf{P}|\mathbf{P}} = \{(\alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_3\beta_1 - \alpha_1\beta_3)^2 + (\alpha_1\beta_2 - \alpha_2\beta_1)^2\}^{\frac{1}{2}}$$
$$= \alpha\beta\left\{1 - \left(\frac{\alpha_1}{\alpha}\frac{\beta_1}{\beta} + \frac{\alpha_2}{\alpha}\frac{\beta_2}{\beta} + \frac{\alpha_3}{\alpha}\frac{\beta_3}{\beta}\right)^2\right\}^{\frac{1}{2}}$$
$$= \alpha\beta \sin\theta.$$

Hence, the magnitude P of the outer product $[\boldsymbol{\alpha}\boldsymbol{\beta}]$ is represented geometrically by the area of the parallelogram constructed upon line-vectors which are the geometrical representations of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. This area together with a unit line-vector normal to it was called by Grassmann a *Plangrösse*, for which we use as an equivalent term Vectorial Area; the direction of the unit line-vector normal is usually chosen by convention so that, if the line-vector representing $\boldsymbol{\alpha}$ undergo a rotation about it toward that representing $\boldsymbol{\beta}$, then the directions of the normal and of the rotation will be related as the thrust and twist of a right-handed screw. With this convention the vectorial area representing the outer product $[\boldsymbol{\beta}\boldsymbol{\alpha}]$ is the negative of that representing the outer product $[\boldsymbol{\alpha}\boldsymbol{\beta}]$, as it should be since $[\boldsymbol{\beta}\boldsymbol{\alpha}] = -[\boldsymbol{\alpha}\boldsymbol{\beta}]$.

Referring again to the expression found above for the outer product of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, viz:

$$\begin{aligned}\mathbf{P} = [\boldsymbol{\alpha}\boldsymbol{\beta}] = \quad &(\alpha_2\beta_3 - \alpha_3\beta_2)\,[\mathbf{e}_2\mathbf{e}_3] \\ &+(\alpha_3\beta_1 - \alpha_1\beta_3)\,[\mathbf{e}_3\mathbf{e}_1] \\ &+(\alpha_1\beta_2 - \alpha_2\beta_1)\,[\mathbf{e}_1\mathbf{e}_2],\end{aligned}$$

the units of the second order on the right represent vectorial areas on the co-ordinate planes, having for their associated unit normals the unit line-vectors representing the primary units $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, respectively; the quantity $(\alpha_2\beta_3 - \alpha_3\beta_2)\,[\mathbf{e}_2\mathbf{e}_3]$ is represented geometrically by a vectorial area in the $\mathbf{e}_2$, $\mathbf{e}_3$-co-ordinate plane whose associated unit normal is $\pm\mathbf{e}_1$, according as the numerical coefficient $\alpha_2\beta_3 - \alpha_3\beta_2$ is positive or negative, and this vectorial area may be considered as the orthogonal projection on the $\mathbf{e}_2$, $\mathbf{e}_3$-co-ordinate plane of that representing $[\boldsymbol{\alpha}\boldsymbol{\beta}]$; the corresponding quantities involving the units $[\mathbf{e}_3\mathbf{e}_1]$ and $[\mathbf{e}_1\mathbf{e}_2]$ each have a similar significance. Consequently, vectorial areas may be resolved into components much in the same manner as line-vectors are resolved into components. It follows that vectorial areas may be added in a manner quite analogous to that in which line-vectors are added.

Another product possessing an important geometrical significance was formed by Grassmann by taking the inner product of an outer product $[\boldsymbol{\alpha}\boldsymbol{\beta}]$ of two primary hyper-numbers $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, with a third primary hyper-number $\boldsymbol{\gamma}$. This product (Q) for 3-dimensional space is evaluated as follows:

$$Q = [\alpha\beta] | \gamma$$
$$= (\alpha_2\beta_3 - \alpha_3\beta_2)\,\gamma_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)\,\gamma_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)\,\gamma_3;$$

hence, expressing the result in determinantal form:

$$Q = \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix}$$

Consequently, Q can be interpreted geometrically as the volume of a parallelepiped constructed upon line-vectors which are the geometrical representations of the primary hyper-numbers α, β, γ, the volume being reckoned positive or negative according as the numerical value of the determinant possesses a $+$ or $-$ sign; this sign will be $+$ or $-$ according as the line-vectors representing α, β, γ, when drawn from a common origin, form a right- or a left-handed system.

Grassmann's inner product of two primary hyper-numbers for the case of 3-dimensional space is equivalent to the negative of the scalar part, and the outer product is equivalent to the vector part, of Hamilton's quaternionic product of two vectors.

While the geometric algebra of Grassmann for 3-dimensional space and the quaternionic algebra of Hamilton have much in common, they differ essentially, as Professor J. Willard Gibbs often and forcibly pointed out, with regard to the status allotted to vectors and vectorial operations. In the theory of quaternions the vector makes its appearance as a subsidiary affiliate of the quaternion, in spite of the fact that the concept of a vector is far simpler than that of a quaternion, while in Grassmann's geometric algebra the vector appears naturally as a basic quantity.

Grassmann must also be given credit as the originator of the idea of treating matrices as hyper-numbers, an idea which was later developed independently by the great English mathematician Arthur Cayley (1821–1895), Sadlerian professor of mathematics at Cambridge. Cayley's great paper *A Memoir on the Theory of Matrices* was published in the *Philosophical Transactions* in 1858; in this paper he showed that the multiplication law for matrices of the same order and with complex elements includes the quaternion as a special case. The importance of matrix theory in the mathematical machinery (including vector and tensor algebra) of modern physics need not be specially emphasized here, but in this connection a prophetic statement made by Cayley's friend, Professor Tait of Edinburgh, is of interest, viz: "Cayley is forging the weapons for future generations of physicists." From the point of view of vector and tensor algebra Cayley's contributions to the

theory of invariants must be considered of outstanding importance; vectors and tensors are, perhaps, most satisfactorily defined through their invariant properties, in accordance, for example, with the procedure followed in Chapter X of the present book.

Chronologically, it would be appropriate to mention here the contributions of the German mathematician Riemann and his followers to the geometric theory of n-dimensional non-Euclidean manifolds, which are basically associated with the development of tensor analysis. This development, however, did not take place until much later, and it therefore seems preferable to defer further reference to these contributions until the historical development of vector analysis for space of 3-dimensions has been more fully traced.

Peter Guthrie Tait (1831–1901), sometime professor of mathematics in Queen's College, later professor of natural philosophy in the University of Edinburgh, attracted to Hamilton's theory of quaternions by its promise of usefulness in physical applications, became, after Hamilton, its chief advocate. From the time when he began the study of quaternions, shortly after his appointment in 1854 to the chair of mathematics in Queen's College, Tait's thought was largely devoted to this subject, and many articles from his pen urged the adoption of quaternions by physicists as the physical calculus *par excellence*, but notwithstanding his able championship the quaternionic cause was doomed to ultimate failure in this direction; quaternions were not precisely what the physicists wanted. In his loyal and enthusiastic support of Hamilton and his quaternionic ideas Tait was not inclined to give proper recognition to Cartesian co-ordinate methods in connection therewith. Apropos of this matter, the following quotation from Hamilton is contained in a letter from Tait to Cayley:[1)] "I regard it as an inelegance, or imperfection in quaternions, or rather in the state to which it has been hitherto unfolded, whenever it becomes or seems to become necessary to have recourse to x, y, z, etc." Cayley could never understand Tait's position in this connection, nor could Tait understand that of Cayley, and many letters relating to the matter passed between them, but their points of view remained irreconcilable. Incidentally, in one of his letters[1)] Cayley made the following significant statement: "I certainly did not get the notion of a matrix in any way through quaternions: it was either directly from that of a determinant, or as a convenient way of expression of the equations:

[1)] The letter is quoted in full in Chapter IV of *The Life and Scientific Work of P. G. Tait*, published in 1911 by the Cambridge University Press.

$$x' = ax + by,$$
$$y' = cx + dy."$$

Professor James Clerk Maxwell (1831–1879), Cavendish professor of physics in Cambridge University, was no doubt greatly influenced in his attitude toward quaternions by his friend Professor Tait. In his celebrated theoretical treatise *Electricity and Magnetism,* published in 1873, he made use in some measure of the notation and terms of quaternions but little of its operational methods, perhaps on account of his fear that they would not be properly understood by the great majority of his readers or, perhaps, because he was not himself fully convinced that quaternionic methods were best adapted to the exposition of his theory.

To Maxwell we owe the significant terms *Convergence* (negative Divergence) and *Curl* of a vector—terms which have become permanent in the nomenclature of mathematical physics and of vector analysis.

While a great admirer of Hamilton's quaternionic system as presented and advocated by Tait, his letters to Tait touching upon the subject indicate that Maxwell himself never became a thoroughgoing quaternionist; in one of these letters[1] he whimsically summarizes some of his difficulties relating to quaternions as follows:

"Here is another question. May one plough with an ox and an ass together? The like of you may write everything and prove everything in 4nions, but in the transition period the bilingual method may help to introduce and explain the more perfect.

"But even when that which is perfect is come that which builds over three axes will be useful for purposes of calculation by the Cassios[2] of the future.

"Now in a bilingual treatise it is troublesome, to say the least, to find that the square of AB is always positive in Cartesians and always negative in 4nions, and that when the thing is mentioned incidentally you do not know what language is being spoken.

"Are the Cartesians to be denied the idea of a vector as a sensible thing in real life till they can recognize in a meter scale one of a peculiar system of square roots of -1?

"It is always awkward when you are discussing, say, kinetic energy, to find that to ensure its being $+$ ve you must stick a $-$ sign to it, and that when you are proving a minimum in certain

[1] The letter is quoted in full in Chapter IV of *The Life and Scientific Work of P. G. Tait,* published in 1911 by the Cambridge University Press.

[2] An allusion to a passage in Othello, Act I, Scene I, l. 18:

"And what was he?
Forsooth, a great arithmetician,"

cases the whole appearance of the proof should be tending toward a maximum.

"What do you recommend for El. and Mag. in such cases?"

In the same letter Maxwell asks of Tait:

"Do you know Grassmann's *Ausdehnungslehre?* Spottiswood spoke of it in Dublin as something above and beyond quaternions. I have not seen it, but Sir William Hamilton of Edinburgh used to say that the greater the extension the smaller the intention."

Unfortunately, Tait's replies to these searching questions of Maxwell's, striking at the very roots of the difficulties which mathematical physicists were finding in the applications of quaternions, are not available.

The question relating to Cartesians appears to indicate a feeling on Maxwell's part that what mathematical physicists really wanted was a vector algebra not divorced from but rather closely associated with Cartesian methods. Had Maxwell himself been acquainted with Grassmann's *Ausdehnungslehre,* it seems probable that he would have been led to adopt a system of vector algebra closely akin to those developed by J. Willard Gibbs in America and Oliver Heaviside in England.

Josiah Willard Gibbs (1839–1903), professor of mathematical physics in Yale College, had printed (1881–84) for private distribution among his students small pamphlets on the *Elements of Vector Analysis.* His viewpoint is set forth in an introductory note as follows:

"The fundamental principles of the following analysis are such as are familiar under a slightly different form to students of quaternions. The manner in which the subject is developed is somewhat different from that followed in treatises on quaternions, being simply to give a suitable notation for those relations between vectors, or between vectors and scalars, which seem most important, and which lend themselves most readily to analytical transformations, and to explain some of these transformations. As a precedent for such a departure from quaternionic usage Clifford's *Kinematics* may be cited. In this connection the name of Grassmann may also be mentioned, to whose system the following method attaches itself in some respects more closely than to that of Hamilton."

Although printed for private circulation, Gibbs's pamphlets on vector analysis became fairly well known generally, and in course of time there arose a considerable controversy on the issue of Quaternions vs. Vector Algebra. On one side were aligned the leading

supporters of quaternions, and on the other Professor Gibbs and Mr. Oliver Heaviside, who stoutly maintained the advantages of a vectorial over a quaternionic treatment of vectors. The question at issue was eventually decided in favor of the advocates of the treatment of vectors by purely vectorial methods.

The vector algebra developed by Gibbs, based on fundamental ideas of both Grassmann and Hamilton, is essentially that now to be found in most textbooks dealing with this subject, and there are indications that even his notation in its most important details may at last receive fairly universal acceptance.

The treatment of the linear vector function with the associated subject of dyadics, a branch of multiple algebra, constitutes without doubt the most original part of Gibbs's contributions to vector analysis. A rather extended treatment of this phase of his work is given in the text of the present book, sufficient it is hoped to enable the reader to obtain a fair idea of the remarkable power and usefulness of the dyadic method. It may be of interest in this connection to mention a remark attributed by Professor Bumstead to Gibbs himself, to the effect that he (Gibbs) had more pleasure in the study of multiple algebra than in any other of his intellectual activities.

The presentation in book form of his work on vector analysis was not undertaken by Gibbs himself, but in 1901 an extensive treatise on the subject based on the lectures of Gibbs was published by Professor E. B. Wilson. This book proved to be of inestimable value in advancing the vector cause.

During the long drawn out controversy over the merits of vectorial and quaternionic methods in physical applications the leading supporter of vectorial methods in England was Oliver Heaviside (1850–1925) who in the early part of his scientific career was a telegraph and telephone engineer. He retired to country life in 1874 and devoted himself to writing, principally on subjects relating to electricity and magnetism. In his book entitled *Electromagnetic Theory*, published in 1893, the tenth chapter, consisting of about 175 pages, is devoted to an exposition of the elements of vectorial algebra and analysis with applications. Here, in the beginning, he makes a vigorous argument in favor of vector methods. The subject itself, with which he had been occupied since 1882, he developed along lines which are quite in harmony with the point of view of Gibbs, although as regards the notation introduced by the latter he frankly states that he does not like it, and adopts one of his own based upon Tait's quaternionic notation. A characteristic

feature of his treatment of the subject is the blending of vectorial with Cartesian methods. The theory of the linear vector function is given by Heaviside in a distinctive form of his own.

The first significant use of vector methods in textbooks on mathematical physics in Germany is found in A. Föppl's *Geometrie der Wirbelfelder* (1897), published as an extension of his earlier book *Einleitung in die Maxwellsche Theorie,* which was rewritten by Max Abraham, and published in two volumes in 1904. The first chapter of the first volume contains a remarkably clear presentation of the algebra and calculus of vectors.

By the beginning of the present century, physicists everywhere were quite convinced that a vector analysis in some such form as that of Gibbs or Heaviside was what they really wanted rather than quaternions. Consequently, textbooks on vector analysis began to make their appearance in America, England, Germany, Italy, and France in numbers too numerous to allow of individual mention here; references to some of them will be found at the end of the present book.

The influence of physical science upon the development of vector analysis during the period which has now been reviewed cannot have escaped the notice of the reader. Without such influence it is quite probable that the subject would have evolved in the hands of pure mathematicians in a manner more logical, perhaps, or at any rate more satisfactory as regards mathematical form. It might, for example, have been developed as a geometric algebra and calculus based upon an invariant theory of orthogonal substitutions along some such lines as those sketched by Professor Felix Klein in his interesting book *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert,* and from the point of view of its applications in modern physics it would, perhaps, have been better had this been the case.

Albert Einstein (1879–1955), while engaged as an engineer in the Swiss patent office, greatly stirred the scientific world by the announcement of his *Restricted Relativity Theory* (1905) in *Wiedermann's Annalen, 17.* In 1914 Einstein accepted a call to the Prussian Academy of Science, Berlin, as successor to the celebrated physical chemist van't Hoff. Two years later his *General Relativity Theory* (1916) was announced in *Wiedermann's Annalen, 49.*

Einstein's revolutionary views on the relativity of physical phenomena created interest of a most lively sort among physicists, philosophers, and mathematicians throughout the world. Mathematicians were specially interested on account of the nature of the

mathematics which Einstein found it expedient and even mandatory to use in the exposition of his theories.

The exposition of the restricted theory, involving the necessity for the discussion of the properties of 4-dimensional pseudo-Euclidean manifolds (space-time), is best made with the aid of vectors and tensors associated with such manifolds; and the exposition of the general theory, involving the discussion of the properties of 4-dimensional non-Euclidean manifolds (space-time), actually demands the use of a special calculus of vectors and tensors associated with such manifolds; fortunately, the calculus required had already been developed by mathematicians in pre-relativity times, but had not then attracted particular notice from physicists. It is essential for the purposes of the present review at least to mention the activities of the various mathematicians which led to the elaboration of this calculus.

A surface represents the only variety of non-Eucidean manifolds capable of actual visualization, and the origin of the metrical geometry of non-Euclidean manifolds in general is found in the work of the German mathematician Karl Friedrich Gauss (1777–1855), professor of mathematics in the University of Göttingen, in which the metrical geometry of an ordinary surface is developed from the standpoint of its intrinsic properties, whereby is meant properties which require for their specification no elements which lie outside the surface itself. Gauss showed that all the metrical properties of any surface figure could be expressed through the coefficients of a differential quadratic form associated with the surface, viz:

$$\overline{ds}^2 = g_{11}du_1du_2 + g_{12}du_1du_2 + g_{21}du_2du_1 + g_{22}du_2du_2, \quad (g_{21} = g_{12}),$$

where $\bar{ds}^2$ represents the square of the distance between two infinitely near points of the surface, and the du's the differentials of co-ordinates intrinsic to the surface and now known as Gaussian co-ordinates (defined in Art. 31).

The metrical properties of surface figures remain undisturbed under any distortion of the surface which does not involve stretching or tearing, and Gauss showed how a curvature (Gaussian) at any point of a surface could be defined, and expressed in terms of the first and second derivatives with respect to the co-ordinates of the g-coefficients of its differential quadratic form, which also has the property of remaining invariant under such distortion.

If new Gaussian co-ordinates for a surface be introduced, the metrical g-coefficients will be altered, but the differential form itself

must be regarded as an invariant, since the distance between the two generic points will not thereby be changed.

Guided to some extent by Gauss's intrinsic geometry of surfaces, Bernhard Riemann (1826–1866), also a professor of mathematics at Göttingen, developed an intrinsic geometry for non-Euclidean manifolds of any number of dimensions. A point in a manifold of n-dimensions is represented by special values assigned to n variable parameters, and the aggregate of all such possible points constitutes the n-dimensional manifold itself, just as the aggregate of the points on a surface constitutes the surface itself. The n variable parameters are called co-ordinates of the manifold. It was Riemann's idea to endow any such manifold with metrical properties by defining the distance between two generic points, whose corresponding co-ordinates differ only by infinitesimal amounts, by making the square of this distance an invariant expressible as a differential quadratic form in the co-ordinate differences of the two points, corresponding to that of Gauss for a surface.

Riemann's differential quadratic form for an n-dimensional manifold is expressed as follows:

$$\overline{ds}^2 = \sum_{i,j} g_{ij} du_i du_j, \quad (g_{ji} = g_{ij}, \quad i,j = 1, 2, \ldots n),$$

where $\overline{ds}^2$ denotes the square of the distance between the two generic points and the du's denote co-ordinate differences for the two points. Following an analogous procedure to that used by Gauss for the case of a surface, Riemann showed how a metrical geometry could be developed with this differential form as a basis for an n-dimensional manifold in which all its metrical properties are determined by the g-coefficients of the form. He also defined what is to be understood by the curvature at any point of such a manifold; the curvature depends upon the first and second derivatives with respect to the co-ordinates of the coefficients of the differential quadratic form which determines the metrical properties of the manifold.

If the coefficients of the differential quadratic form for an n-dimensional manifold are all constants or can all be made so by transformation of the co-ordinates, the manifold is said to be Euclidean. At all points of a Euclidean manifold the Riemannian curvature vanishes.

In his *Pariser Preisarbeit* (1861) Riemann established the conditions under which the fundamental differential quadratic form for an n-dimensional manifold will reduce to a sum of squares of the co-ordinate differentials, in which case the manifold is Euclidean.

Riemann's fundamental ideas relating to non-Euclidean manifolds were set forth at Göttingen in 1854 in his *Habilitationsvortrag,* which was published as an article in the *Göttinger Abhandlungen* in 1868, two years after his death. The publication of this article created intense interest among the mathematicians of the time and some of them set to work to cultivate the promising field which Riemann's original concepts disclosed. Among the most prominent of these were Beltrami, Christoffel, and Lipschitz. The contributions of these mathematicians to the geometry and calculus of Riemannian manifolds were of special importance in connection with the subject matter of the present review in that they prepared the way for the development of an algebra and calculus of such manifolds—the algebra and calculus of tensors.

This development was due chiefly to the great Italian geometer G. Ricci (1853–1925), professor of mathematics in the University of Palermo, who must be regarded as the founder of the tensor calculus, now fully recognized as one of the most important mathematical aids in theoretical physics. A memoir on the *Methodes de calcul differential et leurs applications,* published in the *Math. Ann.,* Vol. 54, 1901, written by Ricci in collaboration with his pupil Levi-Civita, gives a systematic account of Ricci's classical researches in the field now under consideration.

Tullio Levi-Civita (1873–), professor of mathematics in the University of Rome, contributed to the improvement of Ricci's absolute calculus through his introduction of the notion of parallelism in connection with the differential geometry of non-Euclidean manifolds. In his book entitled *Lezioni di calcolo differenziale assoluto* (1925) his concept of parallelism is expounded at length and a presentation is given *in extenso* of the principles of tensor algebra and calculus. An excellent English translation of this book was made by Miss Marjorie Long and published under the title *The Absolute Differential Calculus.*

References to other important texts dealing with the absolute calculus and its relativity applications will be found at the end of the present book.

The present review cannot be brought to a close without a specific reference to a paper entitled *Vektorielle Begründung der Differentialgeometrie* by Gerhard Hessenberg of Breslau, which appeared in 1916 in volume 78 of the *Mathematische Annalen.* In this paper Hessenberg directs attention to the advantages secured by regarding tensors as hyper-numbers in the sense of Grassmann's *Extensive Grössen* (see above). Conformably to Hessenberg's views,

a tensor can be identified with a homogeneous multilinear form in primary base-vectors which is invariant under co-ordinate transformations; the scalar coefficients of this form are commonly known as tensor components. For example, a homogeneous linear form represents a tensor of the first rank, or vector; a homogeneous bi-linear form represents a tensor of the second rank, or dyadic. In co-ordinate transformations the components of a tensor do not behave as invariants, but the tensor itself is an invariant, and to this fact can be traced the chief advantages arising from the consideration of tensors as hyper-numbers in the Grassmannian sense.

The treatment of tensors in the last chapter of the present book is in conformity with the point of view of Hessenberg.

VECTOR AND TENSOR ANALYSIS

CHAPTER I

THE ELEMENTS OF VECTOR ALGEBRA

§1

Scalar and Vector Quantities—Line-Vectors

Among the quantities dealt with by mathematicians and physicists there are many which, for their purposes, are adequately characterized by the specification of their magnitudes[1] only; such, for example, are mass, volume, temperature and entropy. Quantities of this sort are called Scalar Quantities.

They have also continually to deal with many quantities each of which, for their purposes, requires the specification of a magnitude and a direction; such, for example, are displacement, velocity, acceleration, force, and electric field intensity. Quantities of this sort are called Vector Quantities.

The simplest example of a vector quantity is one which in this book will be called a Line-Vector.

If P be any point in space and Q any other point, then a directed straight line-segment from P to Q is called a Line-Vector.

The pure number expressing its length in terms of the unit of length adopted is called the magnitude of a line-vector.

Two line-vectors of the same magnitude and direction are said to be equal.

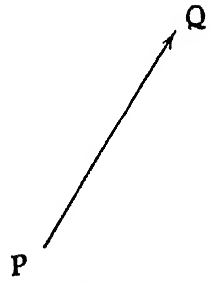


Fig. 1.

Two or more line-vectors are said to be collinear when parallel to the same line, and coplanar when parallel to the same plane.

A line-vector from P to Q can be represented graphically, as in Fig. 1, by an arrow with its tail at the initial point P and its tip at the terminal point Q. It may be designated by $\overrightarrow{PQ}$. For the

[1] The magnitude of any quantity is a pure number expressing the ratio of the size of the quantity to the size of the unit adopted for its measurement.

present a line-vector will usually be denoted by a single letter, **v** for example, in bold-faced type and its length by the same letter, v, in italic type; later, when it becomes necessary to make a distinction between line-vectors and vectors, an underlined letter in bold-faced type will be used to denote a line-vector.

The primary operations of Vector Algebra are concerned with the addition and subtraction of line-vectors in accordance with rules which, as will be seen, are analogues of those governing the corresponding operations of Scalar Algebra.

§2

Addition and Subtraction of Line-Vectors

A line-vector **b** is added to a line-vector **a** by adjoining its initial point to the terminal point of **a** as shown in Fig. 2(a), and then drawing a line-vector from the initial point of **a** to the terminal point of **b**; this line-vector is called the resultant or sum obtained by adding **b** to **a** and is denoted by **a** + **b**.

In a similar manner **a** can be added to **b** as shown in Fig. 2(b), and the resultant or sum denoted by **b** + **a**. It is evident that

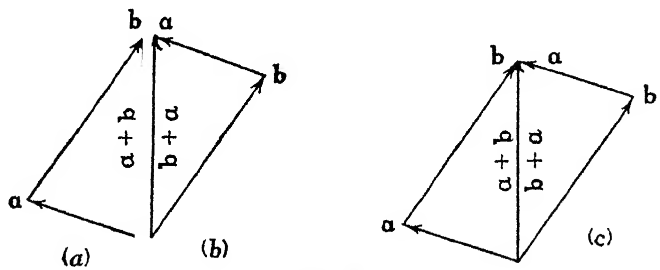


Fig. 2.

figures (a) and (b) are equal triangles which may be adjoined, as shown in Fig. 2(c), to form a parallelogram with a vector diagonal equivalently represented by **a** + **b** or by **b** + **a**. Therefore, the resultant or sum of two line-vectors can be obtained by drawing them from a common origin, completing a parallelogram upon them as sides, and drawing a line-vector from their common origin to the opposite vertex of the parallelogram, this line-vector being the resultant or sum required. The two line-vectors are therefore said to be added in accordance with the "parallelogram law" of addition.

Since $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$, the commutative law of addition is obeyed in the addition of two line-vectors.

To add three line-vectors $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ in the sequence given, let them be adjoined as shown in Fig. 3; then from the initial point of $\mathfrak{a}$ draw a line-vector to the terminal point of $\mathfrak{c}$ to obtain the resultant or sum, denoted by $\mathfrak{a} + \mathfrak{b} + \mathfrak{c}$. From Fig. 3 it is evident that the associative law of addition is valid for three line-vectors, and with the aid of the commutative law of addition for two line-vectors it follows that:

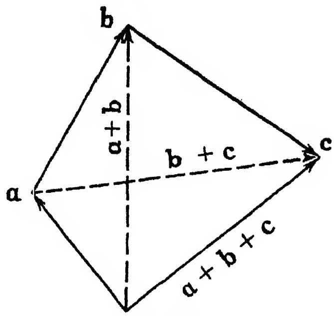


Fig. 3.

$$\mathfrak{a} + \mathfrak{b} + \mathfrak{c} = \mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) = (\mathfrak{b} + \mathfrak{c}) + \mathfrak{a} = \mathfrak{b} + \mathfrak{c} + \mathfrak{a},$$
$$\mathfrak{b} + \mathfrak{c} + \mathfrak{a} = \mathfrak{b} + (\mathfrak{c} + \mathfrak{a}) = (\mathfrak{c} + \mathfrak{a}) + \mathfrak{b} = \mathfrak{c} + \mathfrak{a} + \mathfrak{b}, \text{ etc.}$$

Hence: In the addition of three line-vectors the order in which they are added is immaterial, and the commutative and associative laws are both valid.

By induction it follows that four or more vectors can be added in a similar way, that the order of their addition is immaterial, and that their addition is subject to the commutative and associative laws.

If $\mathbf{v}$ denote any line-vector, then $-\mathbf{v}$ represents (definitionally) a line-vector of the same magnitude as $\mathbf{v}$ but oppositely directed, and $-\mathbf{v}$ is called the negative of $\mathbf{v}$.

Subtraction of a line-vector is the same thing as addition of its negative. For example:

$$\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{v}_1 + (-\mathbf{v}_2).$$

The operation of subtraction of line-vectors requires no further discussion, since it can always be reduced to one of addition.

§3

Multiplication of Line-Vectors by Numbers

If a line-vector be multiplied by the number -1, the product is a line-vector which is the negative of the original vector. The product of a line-vector $\mathbf{v}$ and any real number m is (definitionally) a line-vector $m\mathbf{v}(= \mathbf{v}m)$ whose magnitude is $|m|$ times as great

as that of $\mathbf{v}$ and which is like or oppositely directed according as m is positive or negative.

In the multiplication of line-vectors by pure numbers the distributive law is valid; for example:

$$(m + n)\,\mathbf{v} = m\mathbf{v} + n\mathbf{v},$$
$$m(\mathbf{v}_1 + \mathbf{v}_2) = m\mathbf{v}_1 + m\mathbf{v}_2;$$

the associative law is also valid; for example:

$$m(n\mathbf{v}) = (mn)\,\mathbf{v} = mn\mathbf{v}.$$

These statements are capable of easy geometrical verification.

§4

Resolution of a Line-Vector into Components

If $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ denote any three non-coplanar line-vectors and if x, y, z be positive or negative pure numbers, suitably chosen, then any line-vector $\mathbf{v}$ can be expressed in the form:

$$\mathbf{v} = x\mathbf{a} + y\mathbf{b} + z\mathbf{c}.$$

For it is evident that $\mathbf{v}$ is the line-vector diagonal of a parallelepiped having edges of lengths xa, yb, zc parallel respectively to $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and that the terminal point of the diagonal can be reached from its initial point by taking the steps $x\mathbf{a}$, $y\mathbf{b}$, $z\mathbf{c}$ in any order along the corresponding edges.

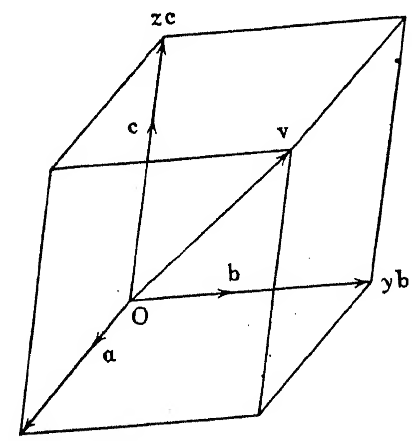


Fig. 4.

The line-vectors $x\mathbf{a}$, $y\mathbf{b}$, $z\mathbf{c}$ are called the Components of $\mathbf{v}$ in the directions of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, respectively. The positive or negative pure numbers x, y, z are called the Measure-Numbers of the respective components.

The line-vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ when drawn from a common origin determine a frame of reference called the $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$-Base-System. The line-vector $\mathbf{v}$ can be represented on this base-system as shown in Fig. 4.

When the base-system is given, it is evident that the measure-numbers for the components of a given line-vector are unique. It follows that two line-vectors referred to the same base-system will be equal if the measure-numbers of their corresponding components are equal in pairs, and vice versa.

If the measure-numbers of the components of a line-vector are all zero, it is called a Null Line-Vector.

§5

Resultant or Sum of any Number of Line-Vectors in Terms of their Components

Let n line-vectors $\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n$ be expressed in terms of their components on an $\mathfrak{a}$, $\mathbf{b}$, $\mathbf{c}$-base-system as follows:

$$\begin{aligned} \mathbf{v}_1 &= x_1\mathfrak{a} + y_1\mathbf{b} + z_1\mathbf{c}, \\ \mathbf{v}_2 &= x_2\mathfrak{a} + y_2\mathbf{b} + z_2\mathbf{c}, \\ &\vdots \\ \mathbf{v}_n &= x_n\mathfrak{a} + y_n\mathbf{b} + z_n\mathbf{c}. \end{aligned}$$

If $\mathbf{v}$ denote the resultant or sum of these line-vectors, then by addition:

$$\begin{aligned} \mathbf{v} &= \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n \\ &= (x_1 + x_2 + \cdots + x_n)\,\mathfrak{a} \\ &\quad + (y_1 + y_2 + \cdots + y_n)\,\mathbf{b} \\ &\quad + (z_1 + z_2 + \cdots + z_n)\,\mathbf{c} \\ &\quad \left(\sum_{s=1}^{n} x_s\right)\mathfrak{a} + \left(\sum_{s=1}^{n} y_s\right)\mathbf{b} + \left(\sum_{s=1}^{n} z_s\right)\mathbf{c}. \end{aligned}$$

Hence: The measure-numbers of the components on a base-system of the resultant or sum of any number of line-vectors are found by adding the measure-numbers of the corresponding components of the individual vectors.

§6

Specification of Vector Quantities—Definition of a Vector

Any vector quantity has two distinct aspects, one of which may be called its Numerical Aspect and the other its Physical Aspect.

In the case of a line-vector the numerical aspect is specified by three pure numbers which just suffice to determine uniquely its magnitude and orientation or direction with respect to a given frame of reference or base-system. These three pure numbers specify the numerical aspect of the line-vector. But a line-vector has also a physical aspect, since it is defined as a directed line-segment, and a line has the physical attribute of length.

If the numerical aspect of a line-vector is known (through the specification of three pure numbers) then, since its physical aspect (a length) is implied in its definition, the specification of the line-vector may be considered complete.

The numerical aspect of any vector quantity, such as velocity, force, momentum, etc., can be specified, as in the case of the line-vector, by three pure numbers, from which its numerical value (magnitude) and its direction with respect to a given base-system can be inferred.

A vector, by definition,[1] *is a set of three pure numbers which specifies uniquely the numerical aspect of a line-vector with respect to a given frame of reference or base-system; its magnitude and direction are those of the line-vector.*

By this definition a vector is a sort of hyper-number,[2] and may be classed as a vector quantity capable of graphical representation by the line-vector whose numerical aspect it specifies.

The definition taken in conjunction with the properties already given to line-vectors tacitly ascribes to vectors the following properties:

(a) Two vectors are equal if the line-vectors representing them are equal.

(b) A unit vector is one whose magnitude is unity and which can therefore be represented by a line-vector of unit length.

(c) A null-vector is one whose magnitude is zero.

(d) The product of a vector by a positive pure number m is a vector capable of representation by a line-vector m times as great as that which represents the original vector.

(e) The sum of two vectors is a vector capable of representation by a line-vector which is the sum of the two line-vectors which represent the two vectors.

[1] The concept of a vector furnished by this definition will be generalized in Chapter IX.

[2] In the sense of Grassmann; as explained in the Historical Introduction.

(f) Equations expressing relations of line-vectors can be replaced by vector equations by substituting for each line-vector the vector which specifies its numerical aspect.

In the interests of brevity we shall often use expessions such as: parallel vectors, perpendicular vectors, coplanar vectors, the angle between two vectors, etc., and, in the light of what has just been said, the meaning of such expressions will be obvious.

In order to bring out more clearly the significance of the concept of a vector as defined above, consider any line-vector $\mathbf{v}$ and let it be expressed in terms of its components on an $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$-base-system as follows:

$$\mathbf{v} = x\mathbf{a} + y\mathbf{b} + z\mathbf{c}.$$

The measure-numbers x, y, z of the components constitute a vector; for, when they are given, the magnitude of $\mathbf{v}$ and the direction of $\mathbf{v}$ with respect to the base-system will be determined.

If θ_1, θ_2, θ_3 denote the angles made by $\mathbf{v}$ with the base-vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, the sets of numbers (x, θ_2, θ_3), (y, θ_3, θ_1), (z, θ_1, θ_2) also constitute vectors, each specifying the numerical aspect of the line-vector $\mathbf{v}$; for each set, when known, determines the magnitude of $\mathbf{v}$ and its direction with respect to the base-system.

Three numbers which constitute a vector on a given base-system will of course become different numbers upon passing to any other base-system, but the latter can be expressed in terms of the former, and vice versa, with the aid of simple rules. These rules, as will be seen later, might be made the basis for the definition of a vector.

The numerical aspect with respect to a given base-system of all vector quantities (whatever their nature) which are of the same magnitude and like-directed is the same, and consequently their numerical aspects can be individually specified by the same vector, and they can be represented graphically by the same line-vector.

In geometrical and physical applications of vector algebra, it is sometimes necessary to deal with a vector quantity which is definitely associated with a given point or line, and in such a case the vector quantity is said to be localized at the point or in the line. A force acting at a point of a non-rigid body is an example of a vector quantity localized at a point; a force acting at a point of a rigid body is an example of a vector quantity localized in a line, for the effect of the force on the motion of the body will not be altered by shifting its point of application along the line of

action of the force. In such cases, of course, the vector quantity will not be fully specified unless the point or line of application is given.

Vector algebra as it will be developed in the following pages is an algebra of pure vectors, that is pure hyper-numbers. That pure vectors are capable of graphical representation by line-vectors is an incidental rather than a fundamental feature of the vector idea; in fact, such graphical representation is not possible for the generalized vectors met with in the later chapters of the book.

Notation. In this book vectors and vector quantities will be denoted by letters printed in bold-faced Roman type, for example: $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots \mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$; and their magnitudes by the corresponding letters $A, B, C, \ldots a, b, c, \ldots$ printed in italic type, or sometimes as follows—$|\mathbf{A}|, |\mathbf{B}|, |\mathbf{C}|, \ldots |\mathbf{a}|, |\mathbf{b}|, |\mathbf{c}|, \ldots$. Line-vectors representing vectors or vector quantities will in future be denoted by letters in bold-faced Roman type which are underlined; for example, if $\mathbf{A}$ denote any vector or vector quantity, then $\underline{\mathbf{A}}$ will denote the line-vector which represents it. A unit vector in the direction of $\mathbf{A}$ will be denoted by $\mathbf{A}_0$ or by $\mathbf{A}/A$.

To avoid possible misunderstanding in the future, it may be mentioned here that from now on, unless otherwise explicitly stated, all equations in which bold-faced characters appear are equations expressing relationships among pure vectors.

§7

The i, j, k-System of Unit Vectors

Let three mutually perpendicular unit line-vectors $\underline{\mathbf{i}}$, $\underline{\mathbf{j}}$, $\underline{\mathbf{k}}$ be drawn from a common origin O to represent three mutually perpendicular unit vectors $(\mathbf{i}, \mathbf{j}, \mathbf{k})$. They must be disposed relatively to one another as in Fig. 5(a) or as in Fig. 5(b). If disposed as in Fig. 5(a), they constitute a right-handed orthogonal system of unit line-vectors, and if disposed as in Fig. 5(b), they constitute a left-handed orthogonal system of unit line-vectors. If the two systems be oriented so that the $\mathbf{i}$-vectors point to the south and the $\mathbf{j}$-vectors to the east, then in the right-handed system the $\mathbf{k}$-vector will point vertically upward and in the left-handed system vertically downward.

If one of the vectors in either system be reversed in direction, the system will thereby be changed from a right to a left or from a left to a right-handed system, as the case may be.

Each system of unit line-vectors may serve to determine a corresponding X, Y, Z-system of rectangular Cartesian co-ordinate axes, one right-handed and the other left-handed, as shown in Fig. 5, (a) and (b).

Whenever subsequently an **i**, **j**, **k**-system of unit vectors or the corresponding rectangular Cartesian system of axes is introduced, it is to be assumed that the system is right-handed unless otherwise stated.

A vector **r** which specifies the numerical aspect of a line-vector drawn from the origin O of a base-system to a point P is called the Position-Vector of P with respect to O.

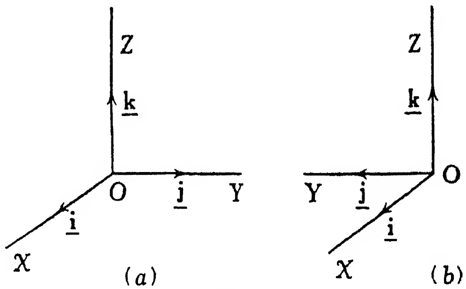
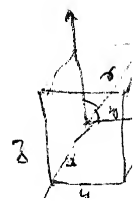


Fig. 5.



If O be the origin of an **i**, **j**, **k**-base-system, we can write:

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \tag{1}$$

where the measure-numbers x, y, z of the components of **r** are the rectangular Cartesian co-ordinates of the point P.

For the magnitude of **r** we have:

$$r = \sqrt{x^2 + y^2 + z^2}; \tag{2}$$

and for the direction cosines of **r** with respect to the axes of the **i**, **j**, **k**-base-system:

$$\cos(\mathbf{r}, \mathbf{i}) = \qquad \cos(\mathbf{r}, \mathbf{j}) = \frac{y}{r}, \qquad \cos(\mathbf{r}, \mathbf{k}) \tag{3}$$

§8

Simple Vector Equations

Suppose a relation to exist among a number of vectors **l**, **m**, **n**, . . . **r**, **s**, **t**, . . . which is expressed by the linear vector equation:

$$\lambda\mathbf{l} + \mu\mathbf{m} + \nu\mathbf{n} + \cdots = \rho\mathbf{r} + \sigma\mathbf{s} + \tau\mathbf{t} + \cdots, \tag{1}$$

where λ, μ, ν, . . . ρ, σ, τ, . . . are numerical coefficients.

Such an equation can always be reduced to one in which the only vectors involved are **i**, **j**, **k**. For any vector can be uniquely resolved into a sum of **i**, **j**, **k**-components and, if we suppose this to have been done for each of the vectors in equation (1), it will take the form:

(2) $$\alpha_1\mathbf{i} + \beta_1\mathbf{j} + \gamma_1\mathbf{k} = \alpha_2\mathbf{i} + \beta_2\mathbf{j} + \gamma_2\mathbf{k},$$

where α_1, β_1, γ_1, α_2, β_2, γ_2 are new coefficients.

The coefficients of like vectors on the two sides of this equation can be equated, since **i**, **j**, **k**, are non-coplanar vectors. Hence:

(3) $$\alpha_1 = \alpha_2, \qquad \beta_1 = \beta_2, \qquad \gamma_1 = \gamma_2.$$

The process of equating coefficients in this manner is permissible, of course, only when the number of vectors involved in the equation is not greater than three, and then only, if these three be non-coplanar. If two vectors only are involved in a vector equation, it is permissible to equate the coefficients of like vectors in the two members of the equation provided the two vectors are non-collinear.

If so desired, all the terms on the right of an equation such as (1) may be transposed with change of sign to the left, giving an equivalent equation with the right-hand member equal to zero.

Geometrical Applications

(a) Vector relations independent of the origin. Suppose $A, B, C, D, E \ldots$ to be a set of points fixed with respect to some frame of reference, and let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e} \ldots$ be the position-vectors of these points with respect to some origin O; furthermore, suppose it has been discovered that some property connected with the points is expressed by the equation:

(1) $$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} + \delta\mathbf{d} + \epsilon\mathbf{e} + \cdots = 0,$$

where $\alpha, \beta, \gamma, \delta, \epsilon \ldots$ are numerical coefficients. Will the property thus expressed be dependent in some measure upon the choice of origin, or will it be quite independent of where the origin is taken? It will be seen that the position of the origin has nothing to do with the matter provided the coefficients satisfy the equation:

(2) $$\alpha + \beta + \gamma + \delta + \epsilon + \cdots = 0.$$

For, let $\mathbf{a}', \mathbf{b}', \mathbf{c}', \mathbf{d}', \mathbf{e}', \ldots$ be the position-vectors of $A, B, C, D, E \ldots$ with respect to a new origin O' whose position-vector with

respect to the origin O is $\mathbf{q}$: then, since $\mathbf{a} = \mathbf{a}' + \mathbf{q}$, $\mathbf{b} = \mathbf{b}' + \mathbf{q}$, . . . , it follows from equation (1) that:

$$\alpha\mathbf{a}' + \beta\mathbf{b}' + \gamma\mathbf{c}' + \delta\mathbf{d}' + \epsilon\mathbf{e}' + \cdots + (\alpha + \beta + \gamma + \delta + \epsilon + \cdots)\mathbf{q} = 0,$$

and hence, if condition (2) is satisfied:

$$\alpha\mathbf{a}' + \beta\mathbf{b}' + \gamma\mathbf{c}' + \delta\mathbf{d}' + \epsilon\mathbf{e}' + \cdots = 0. \tag{3}$$

It is now evident upon comparison of equations (1) and (3) that, if equations (1) and (2) are both satisfied, then (1) will represent some property of the points A, B, C, D, E . . . which is independent of the origin.

(b) To divide a line in a given ratio. Referring to Fig. 6, let AB be a line which it is desired to divide in such a manner that

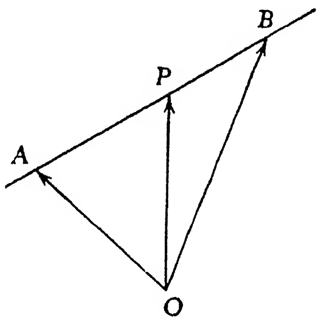


Fig. 6.

$m\overline{AP} = n\overline{PB}$. Let P be the point of division required and let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{p}$ be the position-vectors of the points A, B, P with respect to an arbitrary origin O. We must then have:

$$m(\mathbf{p} - \mathbf{a}) = n(\mathbf{b} - \mathbf{p}),$$

and hence:

$$\mathbf{p} = \frac{m\mathbf{a} + n\mathbf{b}}{m + n}. \tag{4}$$

This equation can be written in the form:

$$(m + n)\mathbf{p} - m\mathbf{a} - n\mathbf{b} = 0,$$

in which the sum of the coefficients of the vectors is zero. The property expressed by equation (4) must therefore be independent of the origin, as is otherwise obvious.

Should the point P divide the line AB externally, then the ratio $m:n$ must be taken with a negative sign prefixed.

(c) Vector equations of a straight line. Let the position in space of the line be specified by requiring it to pass through two points A, B whose position-vectors are $\mathbf{a}$, $\mathbf{b}$ with respect to an arbitrary origin O. See Fig. 7.

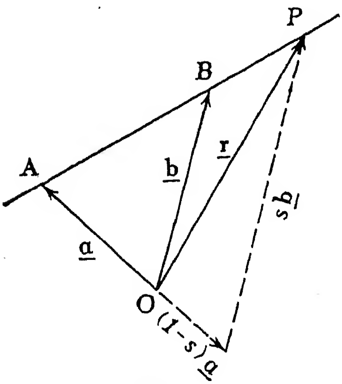


Fig. 7.

Let P be any point on the line (the running point), and let $\mathbf{r}$ be the position-vector of P with respect to O. To obtain the vector equation of the line it is only necessary to find an expression for $\mathbf{r}$ in terms of $\mathbf{a}$ and $\mathbf{b}$ and some scalar variable whose value depends upon the position of P. This is done as follows: Noting that the line-vector $\underline{b} - \underline{a}$ is collinear with the line, it is clear that P can be reached from O by taking successively the steps $\underline{a}$ and $s(\underline{b} - \underline{a})$, provided the value of s is such that $s(\underline{b} - \underline{a}) = \overrightarrow{AP}$; and upon noting further that P may also be reached from O by taking the single step $\underline{r}$, it is evident that the required equation of the line can be written in the form:

$$\mathbf{r} = \mathbf{a} + s(\mathbf{b} - \mathbf{a}),$$

or:

$$\mathbf{r} = (1 - s)\,\mathbf{a} + s\mathbf{b}. \tag{5}$$

The line-vectors representing the two vectors on the right of this equation are indicated by dotted lines in the figure. If $s = 0$, then $\mathbf{r} = \mathbf{a}$; if $s = 1$, then $\mathbf{r} = \mathbf{b}$. To any value of s there will correspond a point on the line, and to each point on the line will correspond a definite value for s. Of course the rôles of $\mathbf{a}$ and $\mathbf{b}$ in this demonstration can be interchanged without introducing anything essentially new.

Writing the equation of the line in the following form:

$$\mathbf{r} - (1 - s)\,\mathbf{a} - s\mathbf{b} = 0,$$

it is seen that the sum of the coefficients of the vectors on the left vanishes, as it should, since the equation simply expresses that the point P is some point on the straight line determined by the points A and B, a property independent of the origin.

A symmetrical form for the equation of the line can be derived as follows: Since the equation must express a relation independent of the origin, assume it to be:

$$x\mathbf{a} + y\mathbf{b} + z\mathbf{r} = 0,$$

with:

$$x + y + z = 0,$$

where x, y, z are variable scalars. Upon elimination of z these two equations give:

$$x(\mathbf{r} - \mathbf{a}) + y(\mathbf{r} - \mathbf{b}) = 0,$$

which shows that $\mathbf{r} - \mathbf{a}$ and $\mathbf{r} - \mathbf{b}$ must be collinear, and hence that $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{r}$ terminate in the straight line through A and B. Solving this equation for $\mathbf{r}$, we find the following symmetrical equation for the line:

$$\mathbf{r} = \frac{x\mathbf{a} + y\mathbf{b}}{x + y}. \tag{6}$$

If $x = 1$, $y = 0$, then $\mathbf{r} = \mathbf{a}$; if $x = 0$, $y = 1$, then $\mathbf{r} = \mathbf{b}$. To any pair of values for x and y except $x = 0$, $y = 0$, there will correspond a point on the line, and the values obtained by multiplying each of these values by any scalar will give another pair of values for x and y corresponding to the same point, as is evident upon inspection of equation (6).

(d) **Vector equations of a plane.** Let the position in space of the plane be specified by requiring it to pass through three points A, B, C whose position-vectors are $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ with respect to an arbitrary origin O. See Fig. 8. Let P be any point on the plane (the running point), and let $\mathbf{r}$ be the position-vector of P with respect to O. Noting that the line-vectors $\underline{\mathbf{b}} - \underline{\mathbf{a}}$, $\underline{\mathbf{c}} - \underline{\mathbf{a}}$ and $\underline{\mathbf{c}} - \underline{\mathbf{b}}$ all lie in the plane, and that any pair of these, say $\underline{\mathbf{b}} - \underline{\mathbf{a}}$ and $\underline{\mathbf{c}} - \underline{\mathbf{a}}$, may be used as a base-system in specifying the position of P in the plane, it is evident that we can write as the equation of the plane:

$$\mathbf{r} = \mathbf{a} + s(\mathbf{b} - \mathbf{a}) + t(\mathbf{c} - \mathbf{a}),$$

or:

$$\mathbf{r} = (1 - s - t)\,\mathbf{a} + s\mathbf{b} + t\mathbf{c}, \tag{7}$$

where s and t are variable scalars. The line-vectors representing the vectors on the right of this equation are indicated by dotted lines in the figure. If $s = t = 0$, then $\mathbf{r} = \mathbf{a}$; if $s = 1$, $t = 0$, then

$\mathbf{r} = \mathbf{b}$; if $s = 0$, $t = 1$, then $\mathbf{r} = \mathbf{c}$. To any pair of values for s and t there will correspond a point on the plane, and vice versa.

The equation of the plane can be written in the form:

$$(1 - s - t)\,\mathbf{a} + s\mathbf{b} + t\mathbf{c} - \mathbf{r} = 0,$$

and it will be noticed that the sum of the scalar coefficients of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{r}$ vanishes, showing that the equation expresses a relation independent of the origin.

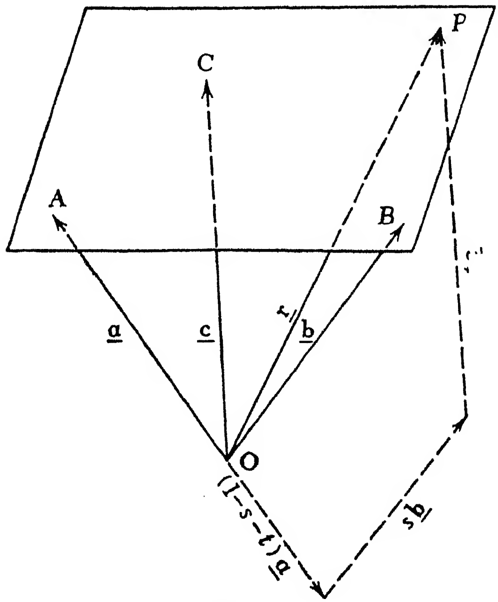


Fig. 8.

The equation of the plane can be derived in a symmetrical form as follows: Since the equation must represent a relation independent of the origin, assume it to be:

$$x\mathbf{a} + y\mathbf{b} + z\mathbf{c} + u\mathbf{r} = 0,$$

with:

$$x + y + z + u = 0,$$

where x, y, z, u are variable scalars. Upon the elimination of u from these two equations we find:

$$x(\mathbf{r} - \mathbf{a}) + y(\mathbf{r} - \mathbf{b}) + z(\mathbf{r} - \mathbf{c}) = 0,$$

which shows that $\mathbf{r} - \mathbf{a}$, $\mathbf{r} - \mathbf{b}$, and $\mathbf{r} - \mathbf{c}$ must be coplanar, and hence that $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{r}$ terminate in the plane through A, B, and C.

Solving this equation for $\mathbf{r}$, we find the following symmetrical equation for the plane:

$$\frac{x\mathbf{a} + y\mathbf{b} + z\mathbf{c}}{x + y + z}. \tag{8}$$

If $x = 1$, $y = z = 0$, then $\mathbf{r} = \mathbf{a}$; if $y = 1$, $z = x = 0$, then $\mathbf{r} = \mathbf{b}$; if $z = 1$, $x = y = 0$, then $\mathbf{r} = \mathbf{c}$. To any set of values for x, y, z, except $x = y = z = 0$, there will correspond a point on the plane, and if each of these values be multiplied by any scalar, a new set of values for x, y, z will be obtained which will also correspond to the same point, as is evident upon inspection of equation (8).

(e) **Vector proof of the proposition that the diagonals of a parallelogram bisect each other.** Referring to Fig. 9, let A, B, C, D

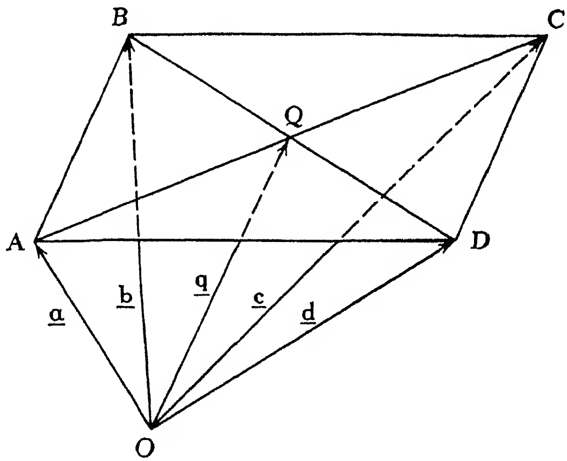


Fig. 9.

be the vertices of the parallelogram, and let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ be the position-vectors of these points with respect to an origin O not in the plane of the figure; also, let Q be the point of intersection of the diagonals, and $\mathbf{q}$ its position-vector with respect to O. Then, by inspection of the figure it is seen that:

$$\begin{aligned} \mathbf{q} &= \mathbf{a} + \alpha(\mathbf{c} - \mathbf{a}), \\ \mathbf{q} &= \mathbf{b} + \beta(\mathbf{d} - \mathbf{b}), \\ \mathbf{q} &= \mathbf{c} + \gamma(\mathbf{a} - \mathbf{c}), \\ \mathbf{q} &= \mathbf{d} + \delta(\mathbf{b} - \mathbf{d}), \end{aligned} \tag{9}$$

where α, β, γ, δ are constants whose values it is now desired to find. Subtracting the second from the first and the fourth from the third of these equations, we get:

$$\mathbf{a} - \mathbf{b} + \alpha(\mathbf{c} - \mathbf{a}) - \beta(\mathbf{d} - \mathbf{b}) = 0,$$
$$\mathbf{c} - \mathbf{d} + \gamma(\mathbf{a} - \mathbf{c}) - \delta(\mathbf{b} - \mathbf{d}) = 0,$$

and upon addition of these two equations, noting that $\mathbf{a} - \mathbf{b} = \mathbf{d} - \mathbf{c}$:

$$(\alpha - \gamma)(\mathbf{c} - \mathbf{a}) + (\delta - \beta)(\mathbf{d} - \mathbf{b}) = 0.$$

Since the vectors $\mathbf{c} - \mathbf{a}$ and $\mathbf{d} - \mathbf{b}$ are not collinear, this equation requires $\alpha = \gamma$, $\beta = \delta$. Upon subtracting the third from the first and the fourth from the second of equations (9), we get:

$$\mathbf{a} - \mathbf{c} - (\alpha + \gamma)(\mathbf{a} - \mathbf{c}) = 0, \qquad \therefore \alpha + \gamma = 1,$$
$$\mathbf{b} - \mathbf{d} - (\beta + \delta)(\mathbf{b} - \mathbf{d}) = 0, \qquad \therefore \beta + \delta = 1.$$

Hence, $\alpha = \beta = \gamma = \delta = 1/2$, and with these values the first and second or the third and fourth of equations (9) show that Q must be the mid-point of each of the diagonals.

Noting that $\alpha = \gamma$ and $\beta = \delta$, we have, by addition of the first and third and of the second and fourth of equations (9):

$$\mathbf{q} = \frac{\mathbf{a} + \mathbf{c}}{2}, \qquad \mathbf{q} = \frac{\mathbf{b} + \mathbf{d}}{2}. \tag{10}$$

From these equations the following symmetrical expression for the position-vector of Q is obtained:

$$\mathbf{q} = \frac{\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}}{4}. \tag{11}$$

The proof would have been somewhat shorter had the origin been chosen at one of the vertices of the parallelogram, but at the cost of loss of symmetry in the equations.

As a general rule in the solution of geometrical problems by vector methods it is advisable to select the origin at random on account of gain in symmetry.

(f) **Vector proof of the proposition that the three median lines joining the middle points of the opposite edges of a tetrahedron meet in a point which bisects each of them.** This problem is introduced in order to exhibit the operation of vector methods in a three dimensional problem.

Referring to Fig. 10, let $ABCD$ be a tetrahedron and L, M, N, U, V, W the mid-points of its edges. Choose an origin O at random, and let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ be the position-vectors of A, B, C, D with respect

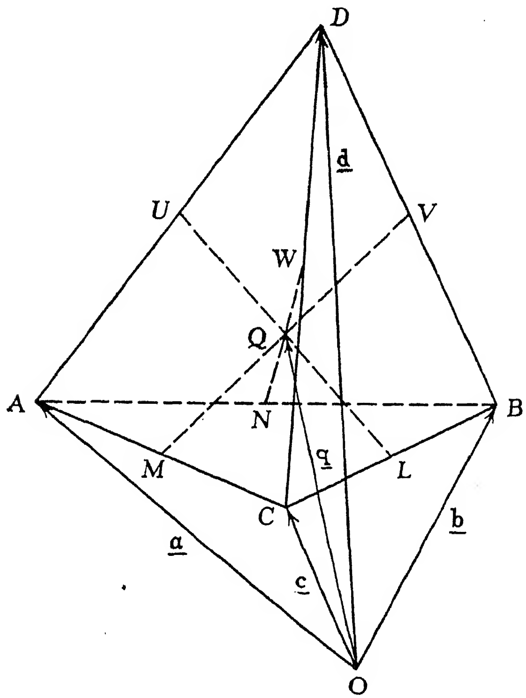


Fig. 10.

to O. Then, if $\mathbf{l}$, $\mathbf{m}$, $\mathbf{n}$, $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are the position-vectors of L, M, N, U, V, W with respect to O:

$$
(12) \qquad \begin{array}{ll}
\mathbf{l} \;:\; \dfrac{\mathbf{b}+\mathbf{c}}{2}, & \mathbf{u} = \dfrac{\mathbf{a}+\mathbf{d}}{2}, \\[2ex]
\;:\; \dfrac{\mathbf{c}+\mathbf{a}}{2}, & \dfrac{\mathbf{b}+\mathbf{d}}{2}, \\[2ex]
\mathbf{a}+\mathbf{b} & \mathbf{w} \quad \mathbf{c}+\mathbf{d}
\end{array}
$$

Now the mid-points of $\overline{LU}$, $\overline{MV}$, $\overline{NW}$ must be, respectively, points of intersection of diagonals of parallelograms on the sides ($\overline{OL}$, $\overline{OU}$),

$(\overline{OM}, \overline{OV})$, $(\overline{ON}, \overline{OW})$, and for the position-vectors of these mid-points we therefore have:

$$\frac{\mathbf{l} + \mathbf{u}}{2} = \frac{\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}}{4},$$

$$\frac{\mathbf{m} + \mathbf{v}}{2} = \frac{\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}}{4},$$

$$\mathbf{n} + \mathbf{w} = \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}$$

Since the right-hand members of these equations are identical, the median lines of the tetrahedron must meet in a common point Q which bisects each of them, and whose position-vector $\mathbf{q}$ with respect to O is given by the equation:

$$\mathbf{q} = \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} \tag{13}$$

§10
Centroids

Let $P_1, P_2, \ldots P_n$ be a system of points having the position-vectors $\mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_n$ with respect to the origin O of an $\mathbf{i}, \mathbf{j}, \mathbf{k}$-system of unit vectors which determines an orthogonal Cartesian system of axes X, Y, Z. If the Centroid C of the System of Points $P_1, P_2, \ldots P_n$ be defined as a point such that its distance from each of the co-ordinate planes is equal to the average distance of all the points from the plane, and if $\mathbf{c}$ be its position-vector with respect to O, the problem of finding $\mathbf{c}$ in terms of $\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_n$ presents itself.

In accordance with the definition of the centroid C its Cartesian co-ordinates c_1, c_2, c_3 will be given by the equations:

$$\begin{aligned} c_1 &= x_1 + x_2 + \quad\quad + x_n \\ c_2 &\quad y_1 + y_2 + \\ c_3 &\quad z_1 + z_2 + \quad\quad + z_n \end{aligned} \tag{1}$$

where the x, y, z's are co-ordinates of $P_1, P_2, \ldots P_n$. Now:

$$\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k},$$
$$\mathbf{r}_1 = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}, \text{ etc.}$$

Hence, upon multiplying equations (1) by **i**, **j**, **k**, respectively, and adding, we obtain for the position-vector of the centroid C:

$$\mathbf{c} = \frac{\mathbf{r}_1 + \mathbf{r}_2 + \cdots + \mathbf{r}_n}{n}. \tag{2}$$

Since there is no reference in this expression to the orientation of Cartesian system of axes used in the definition of the centroid C, it is evident that the position of C is independent of the orientation of that system. Furthermore, the position of C must be independent of the choice of origin; for, upon transposing the vector **c** to the right-hand side of the last equation, we obtain a vector equation with one member zero and the sum of the coefficients of the vectors in the other member equal to zero.

Now suppose masses whose magnitudes are $m_1, m_2 \ldots m_n$ to be concentrated at the points $P_1, P_2 \ldots P_n$. We shall then have a system of mass particles whose position-vectors with respect to the arbitrary origin O are $\mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_n$. The Centroid of the System of Mass-Particles may be defined as a point such that its distance from each of the co-ordinate planes is equal to the sum of the product of each individual mass into its distance from the plane divided by the sum of the masses of all the particles.[1] Let G denote the centroid of the system of mass particles and g_1, g_2, g_3 its Cartesian co-ordinates. Then, in accordance with the definition:

$$\begin{aligned} & \frac{m_1x_1 + m_2x_2 + \cdots + m_nx_n}{m_1 + m_2 + \cdots + m_n}, \\ g_2 &= \frac{m_1y_1 + m_2y_2 + \cdots + m_ny_n}{m_1 + m_2 + \cdots + m_n}, \\ g_3 &= \frac{m_1z_1 + m_2z_2 + \cdots + m_nz_n}{m_1 + m_2 + \cdots + m_n} \end{aligned} \tag{3}$$

Multiplying these equations respectively by **i**, **j**, **k**, and adding, we find for the position-vector **g** of the centroid G of the system of mass particles:

$$\mathbf{g} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + \cdots + m_n\mathbf{r}_n}{m_1 + m_2 + \cdots + m_n}. \tag{4}$$

In this expression no reference to the orientation of system of axes used in the definition of the centroid G is involved, and hence its position must be independent of the orientation of that system of axes. Moreover, the position of G is independent of the choice of origin.

[1] The centroid as here defined is, of course, the center of gravity of the system of mass particles as defined in physics.

For, upon transposing the vector $\mathbf{g}$ to the right-hand side of the last equation, we obtain a vector equation with one member zero and the sum of the coefficients of the vectors in the other member equal to zero.

§11

The Scalar or Direct Product of Two Vectors

The Scalar or Direct Product of two vectors is, by definition, the scalar quantity equal to the product of their magnitudes into the cosine of the angle between their directions.

It is denoted by writing the vectors with a dot between them and is therefore sometimes called the "Dot-Product."

In accordance with the definition, if $\mathfrak{a}$ and $\mathbf{b}$ denote any two vectors:

$$\mathfrak{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathfrak{a} = ab \cos(\mathfrak{a}, \mathbf{b}). \tag{1}$$

Obviously, the scalar product of any two vectors $\mathfrak{a}$, $\mathbf{b}$ is subject to the commutative law of multiplication: $\mathfrak{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathfrak{a}$.

Multiplication of either vector of a scalar product by a scalar is equivalent to multiplication of the product by the scalar.

If two vectors are mutually perpendicular their scalar product vanishes, since the cosine of the angle between their directions is then zero; and conversely.

If $\mathbf{l}$ and $\mathbf{m}$ are unit vectors, then their scalar product $\mathbf{l} \cdot \mathbf{m}$ is equal to the cosine of the angle between their directions.

The scalar product of a vector by itself is equal to the square of its magnitude: $\mathfrak{a} \cdot \mathfrak{a} = a^2$. If $\mathfrak{a} \cdot \mathfrak{a} = 0$, then $\mathfrak{a}$ is a zero- or null-vector.

For the scalar products of the three unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ in pairs we have by equation (1):

$$\begin{aligned} \mathbf{i} \cdot \mathbf{i} &= 1, \qquad & \mathbf{i} \cdot \mathbf{j} &= \mathbf{j} \cdot \mathbf{i} = 0, \\ \mathbf{j} \cdot \mathbf{j} &= 1, & \mathbf{j} \cdot \mathbf{k} &= \mathbf{k} \cdot \mathbf{j} = 0, \\ \mathbf{k} \cdot \mathbf{k} &= 1, & \mathbf{k} \cdot \mathbf{i} &= \mathbf{i} \cdot \mathbf{k} = 0. \end{aligned} \tag{2}$$

The scalar product of two vectors obeys the distributive as well as the commutative law of multiplication. Thus, for example:

$$\mathfrak{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathfrak{a} \cdot \mathbf{b} + \mathfrak{a} \cdot \mathbf{c}.$$

For, the scalar projection of the line-vector $\underline{b} + \underline{c}$ upon any line parallel to $\underline{a}$ must be equal to the sum of the projections of the

line-vectors $\underline{\mathbf{b}}$ and $\underline{\mathbf{c}}$ upon this line. Hence, if a denote the magnitude of $\mathbf{a}$:

$$\frac{\mathbf{a}}{a}\cdot(\mathbf{b}+\mathbf{c}) = \frac{\mathbf{a}}{a}\cdot\mathbf{b} + \frac{\mathbf{a}}{a}\cdot\mathbf{c},$$

or:

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c}. \tag{3}$$

By induction from this result it is easily proved that the distributive law in the scalar multiplication of vectors is valid in general. For example:

$$(\mathbf{a}+\mathbf{b})\cdot(\mathbf{c}+\mathbf{d}) = \mathbf{a}\cdot\mathbf{c} + \mathbf{a}\cdot\mathbf{d} + \mathbf{b}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{d}.$$

Consider two vectors $\mathbf{a}$ and $\mathbf{b}$ expressed in terms of their components parallel to rectangular Cartesian axes:

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k},$$
$$\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k},$$

the subscripts 1, 2, 3, in the coefficients on the right referring to the X, Y, Z axes, respectively. Then:

$$\mathbf{a}\cdot\mathbf{b} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\cdot(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}).$$

Upon expanding the indicated product on the right with the aid of the distributive law and the scalar product relations (2), we find:

$$\mathbf{a}\cdot\mathbf{b} = ab\cos(\mathbf{a},\mathbf{b}) = a_1b_1 + a_2b_2 + a_3b_3. \tag{4}$$

This equation expresses the scalar product of $\mathbf{a}$ and $\mathbf{b}$ as a sum of products of the corresponding measure-numbers of their X, Y, Z-components.

From equation (4), by taking $\mathbf{b}$ equal to $\mathbf{a}$, we get:

$$\mathbf{a}\cdot\mathbf{a} = a^2 = a_1^2 + a_2^2 + a_3^2.$$

Hence, for the magnitude of $\mathbf{a}$ we have:

$$a = \sqrt{a_1^2 + a_2^2 + a_3^2}, \tag{5}$$

and for the direction cosines of $\mathbf{a}$ with respect to the positive directions of the X, Y, Z-axes:

$$\cos(\mathbf{a},\mathbf{i}) = \frac{a_1}{a}, \qquad \cos(\mathbf{a},\mathbf{j}) = \frac{a_2}{a}, \qquad \cos(\mathbf{a},\mathbf{k}) = \frac{a_3}{a}. \tag{6}$$

Again, from equation (4) we have:

$$\cos(\mathbf{a},\mathbf{b}) = \frac{a_1}{a}\frac{b_1}{b} + \frac{a_2}{a}\frac{b_2}{b} + \frac{a_3}{a}\frac{b_3}{b}. \tag{7}$$

This equation expresses the cosine of the angle between the directions of $\mathbf{a}$ and $\mathbf{b}$ as a sum of products of their corresponding direction cosines.

§12
Applications Involving Scalar Products of Two Vectors

The scalar product of two vectors makes its appearance in physical and geometrical problems treated by vector methods whenever the cosine of an angle between two directions comes into consideration.

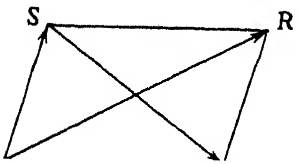


Fig. 11.

(a) Consider the parallelogram $PQRS$ shown in Fig. 11. Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, $\mathbf{d}$ be the vectors represented by the line-vector sides $\overrightarrow{PQ}$, $\overrightarrow{PS}$ and diagonals $\overrightarrow{PR}$, $\overrightarrow{SQ}$. Then $\mathbf{c} = \mathbf{a} + \mathbf{b}$ and $\mathbf{d} = \mathbf{a} - \mathbf{b}$. Hence, with the aid of the distributive law, we have:

$$c^2 = (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) = a^2 + 2ab \cos(\mathbf{a}, \mathbf{b}) + b^2,$$
$$d^2 = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = a^2 - 2ab \cos(\mathbf{a}, \mathbf{b}) + b^2.$$

These are the familiar equations of geometry expressing the squares of the magnitudes of the diagonals of a parallelogram in terms of those of two of its sides and the cosine of the angle between them.

(b) As a second example consider the projection of a plane area

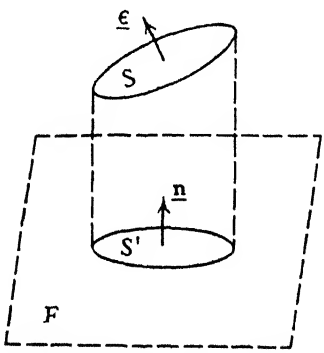


Fig. 12.

S upon a plane F. See Fig. 12. If S' denote the projected

area and $\boldsymbol{\varepsilon}$, $\mathbf{n}$ unit vectors indicating the direction of normals to S and S', then S' is expressed in terms of S by the equation:

$$S' = \pm\boldsymbol{\varepsilon}\cdot\mathbf{n}S,$$

where, since S and S' must be positive, the + or − sign is to be taken according as the angle between $\boldsymbol{\varepsilon}$ and $\mathbf{n}$ is acute or obtuse.

§13

Orthogonal Transformation of an i, j, k-System of Unit Vectors with Origin Fixed

By rotation about the origin O the orthogonal set of line-vectors representing an $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system of unit vectors is transformed into a congruent set representing the transformed system of unit vectors which will be designated the $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$-system. It is assumed that in the transformation $\mathbf{i}$ goes into $\mathbf{i}'$, $\mathbf{j}$ into $\mathbf{j}'$, and $\mathbf{k}$ into $\mathbf{k}'$. The transformation equations expressing $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ in terms of $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ and vice versa are required.

If $\mathbf{r}$ denote the position-vector of any point P with respect to the origin O, then:

$$\mathbf{r} = \mathbf{i}\cdot\mathbf{r}\mathbf{i} + \mathbf{j}\cdot\mathbf{r}\mathbf{j} + \mathbf{k}\cdot\mathbf{r}\mathbf{k}, \tag{1}$$

$$\mathbf{r} = \mathbf{i}'\cdot\mathbf{r}\mathbf{i}' + \mathbf{j}'\cdot\mathbf{r}\mathbf{j}' + \mathbf{k}'\cdot\mathbf{r}\mathbf{k}'. \tag{2}$$

If in equation (1) the vector $\mathbf{r}$ be taken in turn as $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$, and if in equation (2) the vector $\mathbf{r}$ be taken in turn as $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, we obtain the required equations of transformation:

$$\begin{aligned} \mathbf{i}' &= \mathbf{i}\cdot\mathbf{i}'\mathbf{i} + \mathbf{j}\cdot\mathbf{i}'\mathbf{j} + \mathbf{k}\cdot\mathbf{i}'\mathbf{k}, & \mathbf{i} &= \mathbf{i}'\cdot\mathbf{i}\mathbf{i}' + \mathbf{j}'\cdot\mathbf{i}\mathbf{j}' + \mathbf{k}'\cdot\mathbf{i}\mathbf{k}', \\ \mathbf{j}' &= \mathbf{i}\cdot\mathbf{j}'\mathbf{i} + \mathbf{j}\cdot\mathbf{j}'\mathbf{j} + \mathbf{k}\cdot\mathbf{j}'\mathbf{k}, & \mathbf{j} &= \mathbf{i}'\cdot\mathbf{j}\mathbf{i}' + \mathbf{j}'\cdot\mathbf{j}\mathbf{j}' + \mathbf{k}'\cdot\mathbf{j}\mathbf{k}', \\ \mathbf{k}' &= \mathbf{i}\cdot\mathbf{k}'\mathbf{i} + \mathbf{j}\cdot\mathbf{k}'\mathbf{j} + \mathbf{k}\cdot\mathbf{k}'\mathbf{k}, & \mathbf{k} &= \mathbf{i}'\cdot\mathbf{k}\mathbf{i}' + \mathbf{j}'\cdot\mathbf{k}\mathbf{j}' + \mathbf{k}'\cdot\mathbf{k}\mathbf{k}'. \end{aligned} \tag{3}$$

In these equations the dot-product coefficients of the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are direction cosines of $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ with respect to the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system, and those of the unit vectors $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ are direction cosines of $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ with respect to the $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$-system. Since only three data are required to specify the configuration of one system relative to the other, there exist six independent relations among these nine direction cosines. Obviously, there must be six relations such as:

$$(\mathbf{i}\cdot\mathbf{i}')^2 + (\mathbf{j}\cdot\mathbf{i}')^2 + (\mathbf{k}\cdot\mathbf{i}')^2 = 1, \qquad (\mathbf{i}'\cdot\mathbf{i})^2 + (\mathbf{j}'\cdot\mathbf{i})^2 + (\mathbf{k}'\cdot\mathbf{i})^2 = 1; \tag{4}$$

and since the **i**, **j**, **k** and the **i**′, **j**′, **k**′-systems are each orthogonal, there must be six additional relations such as:

$$\text{(5)} \qquad \begin{aligned} \mathbf{i}\cdot\mathbf{i}'\mathbf{i}\cdot\mathbf{j}' + \mathbf{j}\cdot\mathbf{i}'\mathbf{j}\cdot\mathbf{j}' + \mathbf{k}\cdot\mathbf{i}'\mathbf{k}\cdot\mathbf{j}' &= 0,\\ \mathbf{i}'\cdot\mathbf{i}\mathbf{i}'\cdot\mathbf{j} + \mathbf{j}'\cdot\mathbf{i}\mathbf{j}'\cdot\mathbf{j} + \mathbf{k}'\cdot\mathbf{i}\mathbf{k}'\cdot\mathbf{j} &= 0. \end{aligned}$$

But six of these twelve relations among the nine direction cosines must be derivable from the other six, since only six independent relations can exist among them.

§14
The Vector Product of Two Vectors

The Vector Product **V** *of a vector* **A** *into a vector* **B** *is defined as a vector whose magnitude is equal to the product of the magnitudes of* **A** *and* **B** *into the sine of the angle between them, and whose*

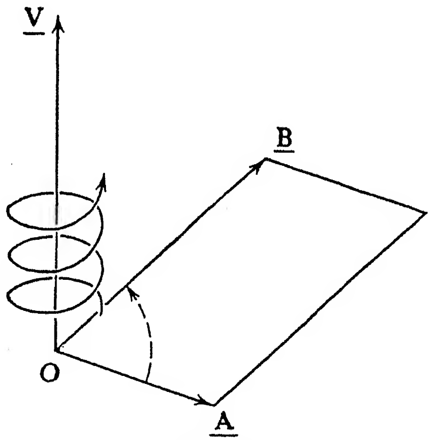


Fig. 13.

direction is specified as follows: if $\underline{A}$, $\underline{B}$, $\underline{V}$ *be line-vectors drawn from a common origin O representing respectively* **A**, **B**, **V**, *and if* $\underline{A}$ *undergo a rotation about* $\underline{V}$ *toward* $\underline{B}$, *then the direction of* $\underline{V}$ *will be related to the direction of the rotation as the thrust to the twist of a right-handed screw.*

The direction relations of **A**, **B**, **V** are indicated by $\underline{A}$, $\underline{B}$, $\underline{V}$ in Fig. 13.

The magnitude of the vector **V** is numerically equal to the area of a parallelogram constructed upon $\underline{A}$ and $\underline{B}$ as sides.

It should be noticed that a change of the unit in terms of which the magnitudes of **A** and **B** are expressed in the ratio $\gamma:1$ would involve a change in the number expressing the magnitude of **A** or **B** in the ratio $1:\gamma$, while the number expressing the magnitude of **V** would be changed in the ratio $1:\gamma^2$.

The vector product of **A** into **B** is denoted by $\mathbf{A}\times\mathbf{B}$ and is therefore often called the "Cross-Product" of **A** into **B.**

If $\boldsymbol{\varepsilon}$ be a unit vector in the direction of $\mathbf{A}\times\mathbf{B}$, then:

(1) $$\mathbf{A}\times\mathbf{B} = \boldsymbol{\varepsilon}AB\sin(\mathbf{A}, \mathbf{B}).$$

From the definition of $\mathbf{A}\times\mathbf{B}$ it follows directly that:

(2) $$\mathbf{B}\times\mathbf{A} = -\mathbf{A}\times\mathbf{B}.$$

The commutative law of multiplication is therefore valid in the vector multiplication of one vector by another, except for sign.

Multiplication of either vector of a vector product by a scalar is equivalent to multiplication of the product by the scalar.

The vector product of two parallel vectors vanishes, since the sine of the angle between them is zero. Conversely, if the vector product of two vectors, neither of which is a null vector, vanishes, then the two vectors must be parallel.

The vector or cross products of the unit vectors **i, j, k** in pairs have, by equation (1), the following values:

(3) $$\begin{aligned} \mathbf{i}\times\mathbf{i} &= 0, & \mathbf{i}\times\mathbf{j} &= -\mathbf{j}\times\mathbf{i} = \mathbf{k},\\ \mathbf{j}\times\mathbf{j} &= 0, & \mathbf{j}\times\mathbf{k} &= -\mathbf{k}\times\mathbf{j} = \mathbf{i},\\ \mathbf{k}\times\mathbf{k} &= 0, & \mathbf{k}\times\mathbf{i} &= -\mathbf{i}\times\mathbf{k} = \mathbf{j}. \end{aligned}$$

The vectors **A** and **B** can be expressed in terms of their components on an **i, j, k**-system of axes as follows:

$$\mathbf{A} = A_1\mathbf{i} + A_2\mathbf{j} + A_3\mathbf{k},$$
$$\mathbf{B} = B_1\mathbf{i} + B_2\mathbf{j} + B_3\mathbf{k},$$

and the vector product $\mathbf{A}\times\mathbf{B}$ can be expressed in terms of its components on the same system as follows.

(4) $$\mathbf{A}\times\mathbf{B} = (A_2B_3 - A_3B_2)\,\mathbf{i} + (A_3B_1 - A_1B_3)\,\mathbf{j} + (A_1B_2 - A_2B_1)\,\mathbf{k},$$

as will now be shown.

Multiplying equation (1) by $\mathbf{i}\cdot$, we get for the X-component of $\mathbf{A}\times\mathbf{B}$:

$$\mathbf{i}\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{i}\cdot\boldsymbol{\varepsilon}AB\sin(\mathbf{A}, \mathbf{B}).$$

The right-hand member of this equation is numerically equal to twice the projected area on the **j**, **k**-plane of the area of the triangle whose vertices are (see Fig. 13) the origin O and the terminal points of $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$. The co-ordinates of these points are respectively (o, o, o), (A_1, A_2, A_3), and (B_1, B_2, B_3). In Fig. 14 the shaded triangle OPQ represents this projected area, and the co-ordinates of its vertices O, P, Q are (o, o, o), (o, A_2, A_3) and (o, B_2, B_3) respectively. This triangle is inscribed in a rectangle whose sides in pairs are numerically equal to A_2 and B_3, and the area of the triangle is equal to the area of the rectangle less the sum of the areas of the three unshaded triangles which are also inscribed in the rectangle. Hence:

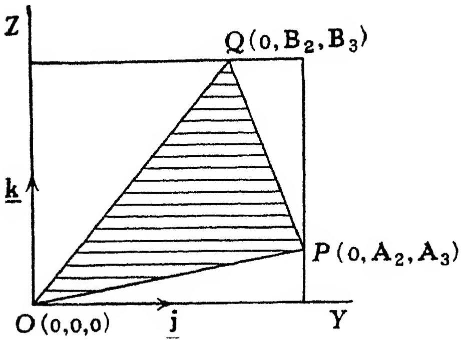


Fig. 14.

$$\mathbf{i} \cdot \varepsilon AB \sin(\mathbf{A}, \mathbf{B}) = 2\left[A_2B_3 - \frac{1}{2}\{A_2A_3 + (B_3 - A_3)(A_2 - B_2) + B_2B_3\}\right]$$
$$= A_2B_3 - A_3B_2.$$

The X-component of $\mathbf{A} \times \mathbf{B}$ is therefore $(A_2B_3 - A_3B_2)\,\mathbf{i}$, and the Y and Z-components can be found by cyclical interchange of the subscripts. The sum of these components must be equal to $\mathbf{A} \times \mathbf{B}$, and equation (4) is therefore valid.

Equation (4) can be written in the determinantal form:

$$(4') \qquad \mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}.$$

With the aid of the relations (3) the magnitude V of $\mathbf{A} \times \mathbf{B}$ can be found in terms of the measure-numbers of its components by squaring both sides of equation (4) and extracting square roots. We thus obtain the equation:

$$(5) \qquad \mathrm{V} = [(A_2B_3 - A_3B_2)^2 + (A_3B_1 - A_1B_3)^2 + (A_1B_2 - A_2B_1)^2]^{\frac{1}{2}}.$$

The direction cosines of $\mathbf{A} \times \mathbf{B}$ with respect to the **i**, **j**, **k**-axes are

$$(6) \qquad \frac{A_2B_3 - A_3B_2}{\mathrm{V}}, \quad \frac{A_3B_1 - A_1B_3}{\mathrm{V}}, \quad \frac{A_1B_2 - A_2B_1}{\mathrm{V}}.$$

If $\mathbf{l}$ and $\mathbf{m}$ denote two unit vectors, they can be expressed in terms of their $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-components as follows:

$$\mathbf{l} = l_1\mathbf{i} + l_2\mathbf{j} + l_3\mathbf{k},$$
$$\mathbf{m} = m_1\mathbf{i} + m_2\mathbf{j} + m_3\mathbf{k},$$

where the measure-numbers of the components of $\mathbf{l}$ and $\mathbf{m}$ respectively are the direction cosines of $\mathbf{l}$ and $\mathbf{m}$. By equations (1) and (4) the following formula for the square of the sine of the angle between the directions of $\mathbf{l}$ and $\mathbf{m}$ is found:

$$\text{(6)} \qquad \sin^2(\mathbf{l}, \mathbf{m}) = (l_2m_3 - l_3m_2)^2 + (l_3m_1 - l_1m_3)^2 + (l_1m_2 - l_2m_1)^2.$$

From equation (4), if $\mathbf{C}+\mathbf{D}$ be written for $\mathbf{B}$, we find:

$$\begin{aligned} \mathbf{A}\times(\mathbf{C}+\mathbf{D}) = \; & [A_2(C_3 + D_3) - A_3(C_2 + D_2)]\,\mathbf{i} \\ & + [A_3(C_1 + D_1) - A_1(C_3 + D_3)]\,\mathbf{j} \\ & + [A_1(C_2 + D_2) - A_2(C_1 + D_1)]\,\mathbf{k} \\ = \; & (A_2C_3 - A_3C_2)\,\mathbf{i} + (A_3C_1 - A_1C_3)\,\mathbf{j} + (A_1C_2 - A_2C_1)\,\mathbf{k} \\ & + (A_2D_3 - A_3D_2)\,\mathbf{i} + (A_3D_1 - A_1D_3)\,\mathbf{j} + (A_1D_2 - A_2D_1)\,\mathbf{k}. \end{aligned}$$

Hence:

$$\text{(7)} \qquad \mathbf{A}\times(\mathbf{C}+\mathbf{D}) = \mathbf{A}\times\mathbf{C} + \mathbf{A}\times\mathbf{D}.$$

The distributive law of multiplication is therefore valid if in the expansion of the indicated product on the left the order of the vectors be maintained. With the aid of this formula the distributive law is easily seen by induction to be valid in the vector multiplication of any compound vector into any other compound vector, provided the order of the vectors be maintained. For example:

$$(\mathbf{A} + \mathbf{B})\times(\mathbf{C} + \mathbf{D}) = \mathbf{A}\times\mathbf{C} + \mathbf{A}\times\mathbf{D} + \mathbf{B}\times\mathbf{C} + \mathbf{B}\times\mathbf{D}.$$

§15

Applications Involving Vector Products of Two Vectors

The vector product of two vectors makes its appearance in geometrical and physical problems treated by vector methods wherever the sine of the angle between two directions comes into consideration. We shall consider several examples by way of illustration.

(a) **Sine of the sum and the difference of two angles.** Let it be required to derive the formulas for the sine of the sum and

of the difference of two angles α and β. Referring to Fig. 15, let $\underline{a}$, $\underline{b}$, $\underline{c}$ represent unit line-vectors in the $\underline{i}$, $\underline{j}$-plane, α being the angle between $\underline{a}$ and $\underline{i}$, β the angle between $\underline{b}$ and $\underline{i}$ and also that between $\underline{c}$ and $\underline{i}$. Then:

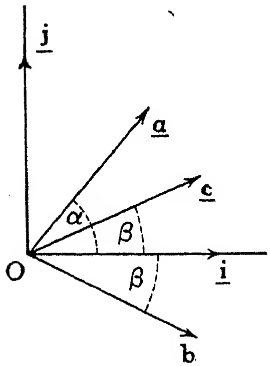


Fig. 15.

$$\mathbf{a} = \cos \alpha \mathbf{i} + \sin \alpha \mathbf{j},$$
$$\mathbf{b} = \cos \beta \mathbf{i} - \sin \beta \mathbf{j},$$
$$\mathbf{c} = \cos \beta \mathbf{i} + \sin \beta \mathbf{j}.$$

From these equations:

$$\mathbf{b} \times \mathbf{a} = (\sin \alpha \cos \beta + \cos \alpha \sin \beta)\, \mathbf{i} \times \mathbf{j},$$
$$\mathbf{c} \times \mathbf{a} = (\sin \alpha \cos \beta - \cos \alpha \sin \beta)\, \mathbf{i} \times \mathbf{j},$$

But, from the figure, and equation (1), Art. 14, we also have:

$$\mathbf{b} \times \mathbf{a} = \sin (\alpha + \beta)\, \mathbf{i} \times \mathbf{j}; \quad \mathbf{c} \times \mathbf{a} = \sin (\alpha - \beta)\, \mathbf{i} \times \mathbf{j}.$$

Consequently:

$$\sin (\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta,$$
$$\sin (\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta.$$

(b) Moment of a localized vector. Let **F** be a vector which specifies the numerical aspect of any vector quantity localized at a point P whose position-vector with respect to an arbitrary origin O

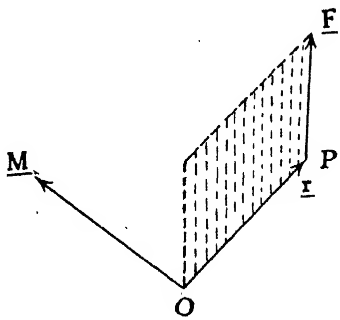


Fig. 16.

is **r**, and let **F** and **r** be represented by line-vectors as shown in Fig. 16.

The vector-moment, **M** say, of the vector **F** with respect to the point O is expressed (conventionally as regards direction) by writing:

(1) $$\mathbf{M} = \mathbf{r} \times \mathbf{F} = \boldsymbol{\varepsilon} rF \sin (\mathbf{r}, \mathbf{F}),$$

where $\boldsymbol{\varepsilon}$ is a unit vector co-directional with the vector product of $\mathbf{r}$ by $\mathbf{F}$, and which is therefore perpendicular to the plane determined by $\mathbf{r}$ and $\mathbf{F}$. Since the vector-moment $\mathbf{M}$ is defined with respect to the point O, it is conventionally represented by a line-vector localized at this point. The magnitude of $\mathbf{M}$ is the product of the magnitude of $\mathbf{F}$ by the factor $r \sin (\mathbf{r}, \mathbf{F})$, which is numerically equal to the perpendicular distance from O to the line through P collinear with $\mathbf{F}$.

If M_1, M_2, M_3 denote the measure-numbers of the components of $\mathbf{M}$ on an orthogonal Cartesian system of axes, F_1, F_2, F_3 those of the components of $\mathbf{F}$, and x, y, z the co-ordinates of P, then by formula (4), Art. 14:

$$
\begin{aligned}
M_1 &= yF_3 - zF_2,\\
M_2 &= zF_1 - xF_3,\\
M_3 &= xF_2 - yF_1.
\end{aligned}
\tag{2}
$$

If $\mathbf{F}$ represent a force acting at the point P of a *rigid* body, it may be considered as localized in the line of action of the force instead of at the point P; for the effect of the force on the motion of the body, by the laws of mechanics, will not be altered by shifting its point of application along this line; it will be noticed that the vector-moment of $\mathbf{F}$ is the same for all points on this line.

(c) Velocity of a point of a rigid body rotating about an axis. Let the vector $\mathbf{v}$ represent the velocity of a point P of a rigid body supposed rotating, with angular velocity specified in magnitude by ω, about an axis. Let $\boldsymbol{\omega}$ be a vector, called the Angular Velocity-Vector, localized in this axis, having a magnitude ω, and such that its direction is related to the direction of rotation as the thrust and twist of a right-handed screw. Choose any point O on the axis as origin, and let $\mathbf{r}$ be the position-vector of P with respect to O. The radius of the circular orbit of P will be equal numerically to $r \sin (\boldsymbol{\omega}, \mathbf{r})$, and the magnitude of the vector $\mathbf{v}$ will be $\omega r \sin (\boldsymbol{\omega}, \mathbf{r})$. The direction of $\mathbf{v}$ will be perpendicular to the plane determined by $\boldsymbol{\omega}$ and $\mathbf{r}$, and, in fact, in the same direction as the vector product of $\boldsymbol{\omega}$ by $\mathbf{r}$. Hence:

$$
\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}. \tag{3}
$$

If the rigid body be subject simultaneously to rotations about a number of axes, each of which passes through the point O, with angular velocities specified by the vectors $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$ . . . , respectively,

and if $\mathbf{v}_1, \mathbf{v}_2 \ldots$ denote the corresponding vector-velocities of P due to the individual rotations, and $\mathbf{v}$ the resultant vector-velocity, then:

$$\begin{aligned}\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots &= \boldsymbol{\omega}_1 \times \mathbf{r} + \boldsymbol{\omega}_2 \times \mathbf{r} + \cdots \\ &= (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2 + \cdots) \times \mathbf{r}.\end{aligned}$$

The body therefore moves as if rotating with an angular velocity specified by the vector:

$$(4) \qquad \boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2 + \cdots .$$

Hence, unlike finite rotations, angular velocities of rotation of a rigid body may be added vectorially.

(d) **Motion of a rigid body with one point fixed.** Referring to Fig. 17, let O be the fixed point, P and Q any two points which lie on the surface of a sphere fixed in the body and with center at O. Suppose that P moves to P' and Q to Q' in a given displacement of the body. Let I be a point of intersection of two great circles on the sphere which, respectively, bisect and are normal to arcs of great circles joining P and P', and Q and Q'. In the spherical triangles IPQ and $IP'Q'$ the sides IP and IP', and the sides IQ and IQ', are equal by construction, and the sides PQ and $P'Q'$ are also equal, since the body is rigid. The two triangles are, therefore, congruent, and in the displacement the original triangle IPQ moves into the position of the triangle $IP'Q'$; the point I and, consequently, all points on the line OI occupy their original positions; the displacement is, therefore, equivalent to a rotation about the line OI. It follows that any infinitesimal displacement of a rigid body moving with one point fixed consists of an infinitesimal rotation about an axis, called the Instantaneous Axis, passing through the fixed point. Hence, taking account of equation (3), the velocity-vector $\mathbf{v}$ of a generic point Q of the rigid body can be expressed as follows:

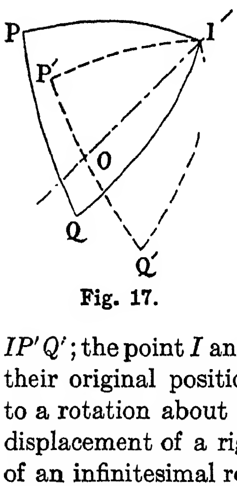


Fig. 17.

$$(5) \qquad \mathbf{v} = \boldsymbol{\omega} \times \mathbf{q},$$

where the vector $\boldsymbol{\omega}$ specifies the angular velocity of rotation of the body about its instantaneous axis, and $\mathbf{q}$ is the position-vector of Q with respect to the fixed point.

§16
The Scalar Triple Product

The scalar product of two vectors one of which is itself the vector product of two vectors is a scalar quantity called a Scalar Triple Product.

Thus, for example, $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, $(\mathbf{a} \times \mathbf{c}) \cdot \mathbf{b}$, and $\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ are scalar triple products of the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$. It is evident that the brackets in these expressions may be removed without ambiguity. For the only interpretation which gives sense to the expression $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$, for example, is that connoted by $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, since $(\mathbf{a} \cdot \mathbf{b}) \times \mathbf{c}$, expressing the vector product of a scalar by a vector, has no meaning.

The scalar triple product $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ can be expressed in terms of the measure-numbers of the components of $\mathbf{a}, \mathbf{b}, \mathbf{c}$ on an $\mathbf{i}, \mathbf{j}, \mathbf{k}$-base-system as follows:

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \quad & a_1(b_2c_3 - b_3c_2) \\ & + a_2(b_3c_1 - b_1c_3) \\ & + a_3(b_1c_2 - b_2c_1), \end{aligned} \tag{1}$$

where the a's, b's, c's are the measure-numbers in question. By inspection it appears that in the expression on the right of equation (1) cyclical interchange of the letters a, b, c is permissible. This implies that cyclical interchange of the vectors on the left is also permissible. Hence:

$$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \mathbf{b} \cdot \mathbf{c} \times \mathbf{a} = \mathbf{c} \cdot \mathbf{a} \times \mathbf{b}.$$

It follows that interchange of $\cdot$ and $\times$ does not alter the value of a scalar triple product. Furthermore, since:

$$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = -\mathbf{a} \cdot \mathbf{c} \times \mathbf{b},$$

it follows that a single non-cyclical interchange of vectors in a scalar triple product does not alter the value of the product, except for change of sign.

If $\underline{a}, \underline{b}, \underline{c}$ are line-vectors representing $\mathbf{a}, \mathbf{b}, \mathbf{c}$, and a parallelepiped be constructed upon them as shown in Fig. 18, and if τ denote the magnitude of its volume, then, as is geometrically evident:

$$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \pm \tau, \quad \begin{cases} + \text{ if } \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ constitute a right-handed system,} \\ - \text{ if } \mathbf{a}, \mathbf{b}, \mathbf{c} \text{ constitute a left-handed system.} \end{cases} \tag{2}$$

The following special cases should be noticed: A scalar triple product will vanish if any two of its vectors are collinear, since

the vector product of two collinear vectors must vanish; the product will also vanish if its three vectors are coplanar, for the triple product in this case will be equal to the scalar product of two perpendicular vectors.

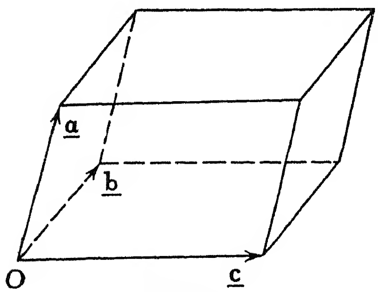


Fig. 18.

If a scalar triple product vanishes, and if no two of its vectors are collinear, and if none of them is a null vector, then the three vectors of the product must be coplanar, since the vector product of any two of them must be perpendicular to the third.

Upon inspection of equation (1) it appears that the scalar triple product $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ can be expressed in determinantal form:

$$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \tag{3}$$

If **A**, **B**, **C** be any three non-coplanar vectors, **a**, **b**, **c** can be expressed in terms of them as follows:

$$\begin{aligned} \mathbf{a} &= A_1\mathbf{A} + A_2\mathbf{B} + A_3\mathbf{C}, \\ \mathbf{b} &= B_1\mathbf{A} + B_2\mathbf{B} + B_3\mathbf{C}, \\ \mathbf{c} &= C_1\mathbf{A} + C_2\mathbf{B} + C_3\mathbf{C}, \end{aligned}$$

where the A's, B's, C's are appropriate measure-numbers. Hence:

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \quad & (A_1\mathbf{A} + A_2\mathbf{B} + A_3\mathbf{C}) \\ & \cdot (B_1\mathbf{A} + B_2\mathbf{B} + B_3\mathbf{C}) \\ & \times (C_1\mathbf{A} + C_2\mathbf{B} + C_3\mathbf{C}). \end{aligned}$$

Upon performing the operations indicated by $\cdot$ and $\times$ in the right-hand member of this equation, we find:

$$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \begin{vmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{vmatrix} \mathbf{A} \cdot \mathbf{B} \times \mathbf{C}. \tag{4}$$

In the special case when **A**, **B**, **C** are taken as **i**, **j**, **k** respectively, this equation becomes equivalent to equation (3).

In what follows, we shall often find it convenient to denote the scalar triple product of three vectors by enclosing them, when written side by side, in square brackets. For example:

$$[\mathbf{abc}] \equiv \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} \equiv \mathbf{b} \cdot \mathbf{c} \times \mathbf{a} \equiv \mathbf{c} \cdot \mathbf{a} \times \mathbf{b}. \tag{5}$$

§17
The Vector Triple Product

The vector product of two vectors one of which is itself the vector product of two vectors is called a Vector Triple Product.

Thus, for example, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ and $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ are vector triple products. It is evident that the parentheses in a vector triple product cannot be removed without ambiguity. For, in general $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is not the same as $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ and, therefore, the associative law is not valid for such products.

Obviously, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is a vector which is perpendicular to both of the vectors $\mathbf{a}$ and $\mathbf{b} \times \mathbf{c}$. It must, therefore, be parallel to the plane determined by $\mathbf{b}$ and $\mathbf{c}$, and consequently expressible in the form:

(1) $$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \beta\mathbf{b} + \gamma\mathbf{c},$$

where β and γ are scalars whose values can easily be found, as will now be shown.

We can express the product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ in terms of its components on an orthogonal Cartesian base-system as follows:

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \; & [a_2(b_1c_2 - b_2c_1) - a_3(b_3c_1 - b_1c_3)]\,\mathbf{i} \\ & + [a_3(b_2c_3 - b_3c_2) - a_1(b_1c_2 - b_2c_1)]\,\mathbf{j} \\ & + [a_1(b_3c_1 - b_1c_3) - a_2(b_2c_3 - b_3c_2)]\,\mathbf{k} \\ = \; & (a_1c_1 + a_2c_2 + a_3c_3)\,(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \\ & - (a_1b_1 + a_2b_2 + a_3b_3)\,(c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}), \end{aligned}$$

where the a's, b's, c's on the right are the measure-numbers of the components of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$. Hence:

(2) $$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{a} \cdot \mathbf{c}\mathbf{b} - \mathbf{a} \cdot \mathbf{b}\mathbf{c}.$$

Upon comparison of equations (1) and (2) we find: $\beta = \mathbf{a} \cdot \mathbf{c}$ and $\gamma = -\mathbf{a} \cdot \mathbf{b}$.

Equation (2) can be written in the form:

(3) $$(\mathbf{b} \times \mathbf{c}) \times \mathbf{a} = \mathbf{a} \cdot \mathbf{b}\mathbf{c} - \mathbf{a} \cdot \mathbf{c}\mathbf{b}.$$

The reduction formulas (2) and (3) are of fundamental importance. With their aid products involving four or more vectors can easily be reduced. For example:

(4) $$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= \mathbf{a} \cdot \mathbf{b} \times (\mathbf{c} \times \mathbf{d}) \\ &= \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{d}\mathbf{c} - \mathbf{b} \cdot \mathbf{c}\mathbf{d}) \\ &= \mathbf{a} \cdot \mathbf{c}\,\mathbf{b} \cdot \mathbf{d} - \mathbf{a} \cdot \mathbf{d}\,\mathbf{b} \cdot \mathbf{c}, \end{aligned}$$

(5) $$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \times \mathbf{b} \cdot \mathbf{d})\,\mathbf{c} - (\mathbf{a} \times \mathbf{b} \cdot \mathbf{c})\,\mathbf{d}.$$

§18

Applications Involving Triple Products

(a) **Equation of a plane.** By means of the scalar triple product the equation of a plane passing through the terminal points of three non-coplanar line-vectors $\underline{\mathrm{a}}$, $\underline{\mathrm{b}}$, $\underline{\mathrm{c}}$ drawn from a common origin O may be simply expressed. Let $\underline{\mathrm{r}}$ be a line-vector from the origin to any point of the plane. See Fig. 19. Then the line-vectors $\underline{\mathrm{r}} - \underline{\mathrm{a}}$, $\underline{\mathrm{b}} - \underline{\mathrm{a}}$, $\underline{\mathrm{c}} - \underline{\mathrm{a}}$ will all lie in the plane. Consequently, the scalar triple product of the vectors $\mathbf{r} - \mathbf{a}$, $\mathbf{b} - \mathbf{a}$, $\mathbf{c} - \mathbf{a}$ which they represent must vanish.

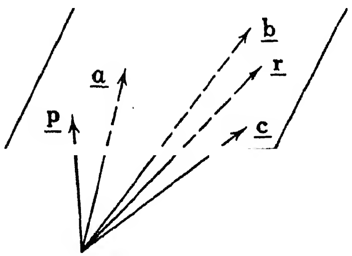


Fig. 19.

We therefore have for the equation of the plane:

$$(1) \qquad (\mathbf{r} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a}) = 0.$$

By expansion of the vector product the equation can be put in the form:

$$(1') \qquad (\mathbf{r} - \mathbf{a}) \cdot (\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}) = 0.$$

It thus appears that $\mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{c} + \mathbf{c} \times \mathbf{a}$ is a vector perpendicular to $\mathbf{r} - \mathbf{a}$ for values of $\mathbf{r}$ corresponding to all points of the plane, and is therefore perpendicular to the plane. Let this vector be denoted by $\mathbf{q}$, and let $\mathbf{p}$ denote a vector perpendicular from the origin upon the plane. To find the value of $\mathbf{p}$, we have from equation (1′):

$$(\mathbf{p} - \mathbf{a}) \cdot \mathbf{q} = 0.$$

But:

$$\mathbf{p} = x\mathbf{q},$$

where x is some scalar. From the last two equations it follows that:

$$(2) \qquad \frac{\mathbf{a} \cdot \mathbf{q}}{\mathbf{q} \cdot \mathbf{q}},$$

$$(3) \qquad \mathbf{p} = \frac{\mathbf{a} \cdot \mathbf{q}}{\mathbf{q} \cdot \mathbf{q}}\,($$

(b) **Reduction of a system of forces to a resultant force and minimum couple.** Any system of forces $\mathbf{F}_1$, $\mathbf{F}_2$, . . . $\mathbf{F}_n$ acting

upon a *rigid* body can be reduced to an equivalent system (as far as the effect upon the motion of the body is concerned) consisting of a single force, called the resultant force, acting through a given point (the point of reduction) and a single couple, called the resultant couple; in particular, by a special choice of the point of reduction the magnitude of the resultant couple may be made a minimum.

Let $\mathbf{f}_1$, $\mathbf{f}_2$ . . . $\mathbf{f}_n$ denote vectors which specify the forces $\mathbf{F}_1$, $\mathbf{F}_2$, . . . $\mathbf{F}_n$ supposed acting at the points P_1, P_2, . . . P_n, respectively, of the rigid body, and let $\mathbf{r}_1, \mathbf{r}_2$ . . . $\mathbf{r}_n$ denote the position-vectors of these points with respect to any arbitrarily chosen point O (the point of reduction). See Fig. 20. The effect of the force $\mathbf{F}_1$ upon the motion of the body is evidently equivalent to that which would be produced by three forces, $\mathbf{F}_1$ acting at P_1, $\mathbf{F}_1$ acting at O, and $-\mathbf{F}_1$ acting at O, if this point be assumed in rigid connection with the body. The force $\mathbf{F}_1$ acting at P_1, and the force $-\mathbf{F}_1$, acting at O, constitute a couple whose vector-moment, by definition, is $\mathbf{r}_1 \times \mathbf{f}_1$. Since each of the forces may be reduced in the same manner, it follows that the original system of forces is equivalent to the forces $\mathbf{F}_1$, $\mathbf{F}_2$, . . . $\mathbf{F}_n$ acting at O, whose vector representatives are $\mathbf{f}_1$, $\mathbf{f}_2$, . . . $\mathbf{f}_n$, together with a system of couples whose individual vector-moments are:

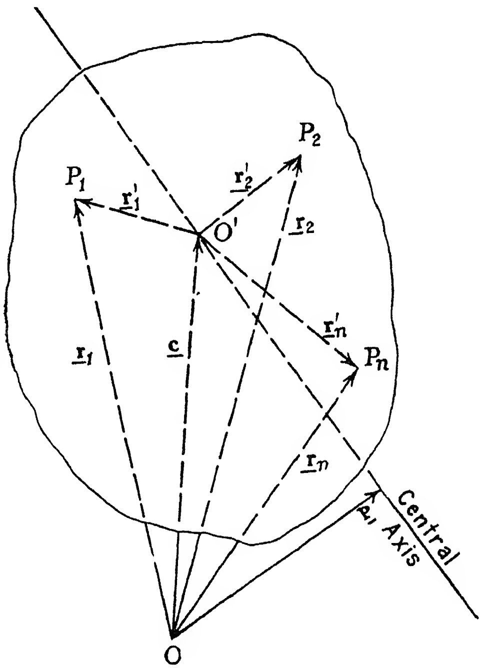


Fig. 20.

$$\mathbf{r}_1 \times \mathbf{f}_1 \qquad \mathbf{r}_2 \times \mathbf{f}_2 \qquad \mathbf{r}_n \times \mathbf{f}_n.$$

As a matter of mechanics it is known that the effect of any number of forces acting at a point of a body is the same as that which would be produced by their resultant acting at the same point, and that the effect of any number of couples acting on a *rigid*

body is the same as that which would be produced by a resultant couple, the vectorial-moment of which is obtained by adding vectorially their individual vectorial-moments. If $\mathbf{R}$ denote a vector specifying the resultant of the individual forces, and $\mathbf{G}$ a vector specifying the moment of the resultant couple for the point of reduction O, then:

$$\mathbf{R} = \mathbf{f}_1 + \mathbf{f}_2 + \cdots + \mathbf{f}_n, \tag{4}$$

$$\mathbf{G} = \mathbf{r}_1 \times \mathbf{f}_1 + \mathbf{r}_2 \times \mathbf{f}_2 + \cdots + \mathbf{r}_2 \times \mathbf{f}_n. \tag{5}$$

We suppose $\mathbf{R}$ and $\mathbf{G}$ not to vanish, and now ask whether by special choice of the point of reduction it is possible to make the direction of $\mathbf{G}$ parallel to the direction of $\mathbf{R}$, which is the same for all points of reduction. Supposing such a point to exist, call it O', and let $\mathbf{c}$ denote the position-vector of O' with respect to O, which we take as origin. The problem now is to find $\mathbf{c}$.

If $\mathbf{G}'$ denote the vector-moment of the resultant couple corresponding to the point of reduction O', then:

$$\mathbf{G}' = \mathbf{r}_1' \times \mathbf{f}_1 + \mathbf{r}_2' \times \mathbf{f}_2 + \cdots + \mathbf{r}_n' \times \mathbf{f}_n,$$

where:

$$\mathbf{r}_1 = \mathbf{r}_1 - \mathbf{c}, \qquad \mathbf{r}_2 = \mathbf{r}_2 - \mathbf{c}, \qquad \cdots \mathbf{r}_n = \mathbf{r}_n - \mathbf{c}.$$

Hence:

$$\begin{array}{l} \mathbf{r}_1 \times \mathbf{f}_1 + \mathbf{r}_2 \times \mathbf{f}_2 + \quad \cdot + \mathbf{r}_n \times \mathbf{f}_n \\ - \mathbf{c} \times \mathbf{f}_1 - \mathbf{c} \times \mathbf{f}_2 - \cdot \quad - \mathbf{c} \times \mathbf{f}_n \end{array}$$

or:

$$\mathbf{G}' = \quad \mathbf{G} + \mathbf{c} \times (-\mathbf{R}). \tag{6}$$

Now, if $\mathbf{G}'$ is to be parallel to $\mathbf{R}$, we must have:

$$\mathbf{R} \times \mathbf{G} + \mathbf{R} \times [\mathbf{c} \times (-\mathbf{R})] = 0,$$

or, upon expansion of the triple vector product:

$$\mathbf{R} \times \mathbf{G} + \mathbf{R} \cdot \mathbf{c}\mathbf{R} - \mathbf{R} \cdot \mathbf{R}\mathbf{c} = 0:$$

This is a linear vector equation in the unknown vector $\mathbf{c}$. A general method for the solution of such equations will be found in Art. 22. Here, we proceed as follows: Multiply both sides of the equation by $\cdot\, \mathbf{R} \times \mathbf{G}$ to obtain:

$$\mathbf{R} \times \mathbf{G} \cdot \mathbf{R} \times \mathbf{G} - \mathbf{R} \cdot \mathbf{R}\mathbf{c} \cdot \mathbf{R} \times \mathbf{G} = 0,$$

which is equivalent to:

$$\left(\mathbf{c} - \frac{\mathbf{R} \times \mathbf{G}}{\mathbf{R} \cdot \mathbf{R}}\right) \cdot \frac{\mathbf{R} \times \mathbf{G}}{\mathbf{R} \cdot \mathbf{R}} = 0. \tag{7}$$

This is the equation of the straight line perpendicular to a vector $\mathbf{p}$, equal to $\mathbf{R} \times \mathbf{G}/\mathbf{R} \cdot \mathbf{R}$ and passing through the terminal point of the line-vector $\underline{\mathbf{p}}$ drawn from O to represent this vector, $\mathbf{c}$ being the running vector for the line. Therefore, any point O' on this line will be a point of reduction such as to make $\mathbf{G}'$ parallel to $\mathbf{R}$. This line is called the Central Axis of the system of forces.

The magnitude of the vector-moment of the resultant couple for a point of reduction which lies on the central axis must be a minimum. This is proved as follows: Multiply equation (6) by $\mathbf{R}\cdot$ to obtain:

$$\mathbf{R} \cdot \mathbf{G}' = \mathbf{R} \cdot \mathbf{G}.$$

This equation shows that $\mathbf{R} \cdot \mathbf{G}$ is invariant to the point of reduction. Since $\mathbf{G}'$ is parallel to $\mathbf{R}$, it follows that the magnitude of $\mathbf{G}$ for any point of reduction not on the central axis must be greater than for a point of reduction on the axis.

§19

Reciprocal Systems of Vectors

If $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ denote any three non-coplanar vectors, and $\mathfrak{a}^1$, $\mathfrak{a}^2$, $\mathfrak{a}^3$ three other vectors defined in terms of $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ by the equations[1]:

$$(1) \qquad \mathfrak{a}^1 = \frac{\mathfrak{a}_2 \times \mathfrak{a}_3}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]}, \qquad \mathfrak{a}^2 = \frac{\mathfrak{a}_3 \times \mathfrak{a}_1}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]}, \qquad \mathfrak{a}^3 = \frac{\mathfrak{a}_1 \times \mathfrak{a}_2}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]},$$

then the systems of vectors $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)$ and $(\mathfrak{a}^1, \mathfrak{a}^2, \mathfrak{a}^3)$ are called Reciprocal Systems. Evidently, the vectors $\mathfrak{a}^1$, $\mathfrak{a}^2$, $\mathfrak{a}^3$ are respectively perpendicular to the planes determined by the pairs of vectors $(\mathfrak{a}_2, \mathfrak{a}_3)$, $(\mathfrak{a}_3, \mathfrak{a}_1)$, $(\mathfrak{a}_1, \mathfrak{a}_2)$.

Upon forming the scalar products of the above expressions for $\mathfrak{a}^1$, $\mathfrak{a}^2$, $\mathfrak{a}^3$ in all possible combinations with $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$, and remembering that a scalar triple product vanishes when it contains the same vector twice, we find:

$$(2) \qquad \begin{aligned} \mathfrak{a}^1 \cdot \mathfrak{a}_1 &= \mathfrak{a}^2 \cdot \mathfrak{a}_2 = \mathfrak{a}^3 \cdot \mathfrak{a}_3 = 1, \\ \mathfrak{a}^2 \cdot \mathfrak{a}_1 &= \mathfrak{a}^3 \cdot \mathfrak{a}_2 = \mathfrak{a}^1 \cdot \mathfrak{a}_3 = 0, \\ \mathfrak{a}^3 \cdot \mathfrak{a}_1 &= \mathfrak{a}^1 \cdot \mathfrak{a}_2 = \mathfrak{a}^2 \cdot \mathfrak{a}_3 = 0. \end{aligned}$$

From the last six of these equations it appears that $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ are respectively perpendicular to the planes determined by the pairs

[1] The justification for the practice here introduced for the first time of using superscripts as identifying indices will appear later.

of vectors $(\mathbf{a}^2, \mathbf{a}^3)$, $(\mathbf{a}^3 \cdot \mathbf{a}^1)$, $(\mathbf{a}^1, \mathbf{a}^2)$, and are therefore respectively proportional to $\mathbf{a}^2 \times \mathbf{a}^3$, $\mathbf{a}^3 \times \mathbf{a}^1$, $\mathbf{a}^1 \times \mathbf{a}^2$; consequently:

$$\mathbf{a}_1 = k_1 \mathbf{a}^2 \times \mathbf{a}^3, \qquad \mathbf{a}_2 = k_2 \mathbf{a}^3 \times \mathbf{a}^1 \qquad \mathbf{a}_3 = k_3 \mathbf{a}^1 \times \mathbf{a}^2,$$

where k_1, k_2, k_3 are factors of proportionality. Upon forming the scalar product of these expressions with the vectors $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$, respectively, and taking account of the first three of equations (2), we obtain:

$$k_1 = k_2 = k_3 = \frac{1}{[\mathbf{a}^1 \mathbf{a}^2 \mathbf{a}^3]}.$$

The vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ can therefore be expressed in terms of the vectors $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ as follows:

$$\mathbf{a}_1 = \frac{\mathbf{a}^2 \times \mathbf{a}^3}{[\mathbf{a}^1 \mathbf{a}^2 \mathbf{a}^3]}, \qquad \mathbf{a}_2 = \frac{\mathbf{a}^3 \times \mathbf{a}^1}{[\mathbf{a}^1 \mathbf{a}^2 \mathbf{a}^3]}, \qquad \mathbf{a}_3 = \frac{\mathbf{a}^1 \times \mathbf{a}^2}{[\mathbf{a}^1 \mathbf{a}^2 \mathbf{a}^3]}. \tag{3}$$

It is on account of the reciprocal nature of the relationships expressed by equations (1) and (3) that the systems of vectors $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ and $(\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3)$ have been called Reciprocal Systems.

It can easily be shown that equations (2) are sufficient as well as necessary conditions that $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ and $(\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3)$ shall be reciprocal systems of vectors. For, let $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$ be any three vectors which satisfy these equations when substituted for $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ respectively. Since $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ are non-coplanar vectors, $\mathbf{A}^1$, $\mathbf{A}^2$, $\mathbf{A}^3$ can be expressed in terms of them as follows:

$$\begin{aligned} \mathbf{A}^1 &= \alpha_{11}\mathbf{a}^1 + \alpha_{12}\mathbf{a}^2 + \alpha_{13}\mathbf{a}^3, \\ \mathbf{A}^2 &= \alpha_{21}\mathbf{a}^1 + \alpha_{22}\mathbf{a}^2 + \alpha_{23}\mathbf{a}^3, \\ \mathbf{A}^3 &= \alpha_{31}\mathbf{a}^1 + \alpha_{32}\mathbf{a}^2 + \alpha_{33}\mathbf{a}^3, \end{aligned} \tag{4}$$

where the α's are appropriate scalar coefficients. Upon forming the scalar products of these expressions in all possible combinations with $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, and taking account of equations (2), we find:

$$\begin{aligned} &\alpha_{11} = \mathbf{a}_1 \cdot \mathbf{A}^1 = 1, && \alpha_{12} = \mathbf{a}_2 \cdot \mathbf{A}^1 = 0, && \alpha_{13} = \mathbf{a}_3 \cdot \mathbf{A}^1 = 0, \\ &\alpha_{21} = \mathbf{a}_1 \cdot \mathbf{A}^2 = 0, && \alpha_{22} = \mathbf{a}_2 \cdot \mathbf{A}^2 = 1, && \alpha_{23} = \mathbf{a}_3 \cdot \mathbf{A}^2 = 0, \\ &\alpha_{31} = \mathbf{a}_1 \cdot \mathbf{A}^3 = 0, && \alpha_{32} = \mathbf{a}_2 \cdot \mathbf{A}^3 = 0, && \alpha_{33} = \mathbf{a}_3 \cdot \mathbf{A}^3 = 1. \end{aligned}$$

It follows then from equations (4) that $\mathbf{A}^1 = \mathbf{a}^1$, $\mathbf{A}^2 = \mathbf{a}^2$, $\mathbf{A}^3 = \mathbf{a}^3$. Hence, the system of vectors $(\mathbf{A}^1, \mathbf{A}^2, \mathbf{A}^3)$ must be reciprocal to the system $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$.

The reciprocal system to the **i**, **j**, **k**-system of unit vectors is the **i**, **j**, **k**-system itself, as is seen at once by replacing $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ by **i**, **j**, **k** in the defining equations (1) for $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$. Conversely,

if $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)$ and $(\mathfrak{a}^1, \mathfrak{a}^2, \mathfrak{a}^3)$ are self-reciprocal systems it follows from equations (2) that each system must constitute either a right-handed or a left-handed orthogonal system of unit vectors.

A relationship of considerable importance between the reciprocal systems of vectors $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)$ and $(\mathfrak{a}^1, \mathfrak{a}^2, \mathfrak{a}^3)$ is that the corresponding scalar triple products $[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]$ and $[\mathfrak{a}^1\mathfrak{a}^2\mathfrak{a}^3]$ are reciprocally related. This is proved as follows:

$$\mathfrak{a}^1 \cdot \mathfrak{a}^2 \times \mathfrak{a}^3 = \frac{(\mathfrak{a}_2 \times \mathfrak{a}_3) \cdot (\mathfrak{a}_3 \times \mathfrak{a}_1) \times (\mathfrak{a}_1 \times \mathfrak{a}_2)}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]^3}.$$

but:

$$\begin{aligned}(\mathfrak{a}_2 \times \mathfrak{a}_3) \cdot (\mathfrak{a}_3 \times \mathfrak{a}_1) \times (\mathfrak{a}_1 \times \mathfrak{a}_2) &= (\mathfrak{a}_2 \times \mathfrak{a}_3) \cdot [(\mathfrak{a}_3 \times \mathfrak{a}_1) \cdot \mathfrak{a}_2\mathfrak{a}_1 \\ &\qquad - (\mathfrak{a}_3 \times \mathfrak{a}_1) \cdot \mathfrak{a}_1\mathfrak{a}_2] \\ &= (\mathfrak{a}_2 \times \mathfrak{a}_3) \cdot \mathfrak{a}_1(\mathfrak{a}_3 \times \mathfrak{a}_1) \cdot \mathfrak{a}_2 \\ &= (\mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\,(\mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3) \\ &= [\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]^2;\end{aligned}$$

therefore:

$$[\mathfrak{a}^1\mathfrak{a}^2\mathfrak{a}^3]\,[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3] = 1. \tag{5}$$

§20
Reciprocal Base-Systems

The two reciprocal systems of vectors $(\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3)$ and $(\mathfrak{a}^1, \mathfrak{a}^2, \mathfrak{a}^3)$ serve to determine two base-systems which are called Reciprocal Base-Systems.

Any vector $\mathfrak{v}$ can be referred to both base-systems by writing:

$$\mathfrak{v} = v^1\mathfrak{a}_1 + v^2\mathfrak{a}_2 + v^3\mathfrak{a}_3, \tag{1}$$

$$\mathfrak{v} = v_1\mathfrak{a}^1 + v_2\mathfrak{a}^2 + v_3\mathfrak{a}^3, \tag{2}$$

where the v's are the measure-numbers of the components of $\mathfrak{v}$. Making use of the properties (equations (2), Art. 19) of vectors belonging to reciprocal systems, the measure-numbers in these equations can be expressed as follows:

$$\begin{aligned} v^1 &= \mathfrak{a}^1 \cdot \mathfrak{v}, & v^2 &= \mathfrak{a}^2 \cdot \mathfrak{v}, & v^3 &= \mathfrak{a}^3 \cdot \mathfrak{v}, \\ v_1 &= \mathfrak{a}_1 \cdot \mathfrak{v}, & v_2 &= \mathfrak{a}_2 \cdot \mathfrak{v}, & v_3 &= \mathfrak{a}_3 \cdot \mathfrak{v}. \end{aligned} \tag{3}$$

Consequently, equations (1) and (2) can be written:

$$\mathfrak{v} = \mathfrak{a}^1 \cdot \mathfrak{v}\mathfrak{a}_1 + \mathfrak{a}^2 \cdot \mathfrak{v}\mathfrak{a}_2 + \mathfrak{a}^3 \cdot \mathfrak{v}\mathfrak{a}_3, \tag{1'}$$

$$\mathfrak{v} = \mathfrak{a}_1 \cdot \mathfrak{v}\mathfrak{a}^1 + \mathfrak{a}_2 \cdot \mathfrak{v}\mathfrak{a}^2 + \mathfrak{a}_3 \cdot \mathfrak{v}\mathfrak{a}^3. \tag{2'}$$

Either base-system or both may be used in connection with the vectorial treatment of a given problem. Suppose, for example, that the work done by a constant force in a displacement is a matter for discussion. The work in question is expressed numerically by $\mathbf{f} \cdot \mathbf{s}$, if $\mathbf{f}$, $\mathbf{s}$ are vectors specifying the force and displacement. Expressing $\mathbf{f}$ and $\mathbf{s}$ in terms of their components on the two base-systems, we have:

$$\mathbf{f} = f^1\mathbf{a}_1 + f^2\mathbf{a}_2 + f^3\mathbf{a}_3, \qquad \mathbf{f} = f_1\mathbf{a}^1 + f_2\mathbf{a}^2 + f_3\mathbf{a}^3,$$
$$\mathbf{s} = s^1\mathbf{a}_1 + s^2\mathbf{a}_2 + s^3\mathbf{a}_3, \qquad \mathbf{s} = s_1\mathbf{a}^1 + s_2\mathbf{a}^2 + s_3\mathbf{a}^3.$$

If the first forms for $\mathbf{f}$ and $\mathbf{s}$ be selected, then:

$$\begin{aligned}\mathbf{f} \cdot \mathbf{s} = \quad & f^1s^1\mathbf{a}_1 \cdot \mathbf{a}_1 + f^1s^2\mathbf{a}_1 \cdot \mathbf{a}_2 + f^1s^3\mathbf{a}_1 \cdot \mathbf{a}_3 \\ & + f^2s^1\mathbf{a}_2 \cdot \mathbf{a}_1 + f^2s^2\mathbf{a}_2 \cdot \mathbf{a}_2 + f^2s^3\mathbf{a}_2 \cdot \mathbf{a}_3 \\ & + f^3s^1\mathbf{a}_3 \cdot \mathbf{a}_1 + f^3s^2\mathbf{a}_3 \cdot \mathbf{a}_2 + f^3s^3\mathbf{a}_3 \cdot \mathbf{a}_3;\end{aligned}$$

and a corresponding form with nine terms for $\mathbf{f} \cdot \mathbf{s}$ would be obtained if the second forms for $\mathbf{f}$ and $\mathbf{s}$ be selected. But, if we select, let us say, the first form for $\mathbf{f}$ and the second form for $\mathbf{s}$, then:

$$\mathbf{f} \cdot \mathbf{s} = f^1s_1 + f^2s_2 + f^3s_3,$$

and a corresponding simple form with three terms would be obtained in case we select the second form for $\mathbf{f}$ and the first form for $\mathbf{s}$.

This example suggests that cases may arise when oblique axes are used in which it may be advantageous to have at disposal two reciprocal base-systems.

§21

Scalar Equations of the First Degree in an Unknown Vector

A scalar equation involving but a single unknown vector, and that not more than once in each of its terms, is a scalar equation of the first degree in the unknown vector; we suppose the unknown vector to be the only unknown quantity in the equation.

Such an equation in its various terms may contain the unknown vector in a large variety of ways, but in all cases the equation may be reduced to a form expressing that the scalar product of the unknown vector and a known vector is equal to a known scalar quantity. It will suffice to illustrate the process whereby this may be done by a single example.

Consider the equation:

(1) $$\alpha\mathbf{a} \cdot \mathbf{r} + \beta\mathbf{b} \cdot \mathbf{c} \times \mathbf{r} + \delta(\mathbf{d} \times \mathbf{e}) \cdot (\mathbf{f} \times \mathbf{r}) + m$$

where all quantities are supposed known except the vector $\mathbf{r}$. We have:

$$\mathbf{b} \cdot \mathbf{c} \times \mathbf{r} = \mathbf{b} \times \mathbf{c} \cdot \mathbf{r},$$
$$(\mathbf{d} \times \mathbf{e}) \cdot (\mathbf{f} \times \mathbf{r}) = (\mathbf{d} \times \mathbf{e}) \times \mathbf{f} \cdot \mathbf{r}.$$

Hence, the equation in question can be written:

$$[\alpha \mathbf{a} + \beta \mathbf{b} \times \mathbf{c} + \delta(\mathbf{d} \times \mathbf{e}) \times \mathbf{f}] \cdot \mathbf{r} = n - m,$$

or, upon writing $\mathbf{u}$ for the known vector expression in square brackets and l for the known scalar $n - m$:

$$(2) \qquad \mathbf{u} \cdot \mathbf{r} = l.$$

This equation does not, of course, determine a unique value for $\mathbf{r}$; in fact, equation (2) is the equation of a plane perpendicular to $\mathbf{u}$ for which $\mathbf{r}$ is the running vector.

A unique value for $\mathbf{r}$ will be determined by three scalar equations, such as:

$$\mathbf{u}_1 \cdot \mathbf{r} = l_1, \quad \mathbf{u}_2 \cdot \mathbf{r} = l_2, \quad \mathbf{u}_3 \cdot \mathbf{r} = l_3,$$

where $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ are known non-coplanar vectors and l_1, l_2, l_3 are known scalars. For, by equation (2′), Art 20, we can write:

$$\mathbf{r} = \mathbf{u}_1 \cdot \mathbf{r}\mathbf{u}^1 + \mathbf{u}_2 \cdot \mathbf{r}\mathbf{u}^2 + \mathbf{u}_3 \cdot \mathbf{r}\mathbf{u}^3,$$

where $\mathbf{u}^1$, $\mathbf{u}^2$, $\mathbf{u}^3$ is the reciprocal system to $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$. Hence:

$$(3) \qquad \mathbf{r} = l_1\mathbf{u}^1 + l_2\mathbf{u}^2 + l_3\mathbf{u}^3,$$

an equation which determines a unique value for $\mathbf{r}$.

Vector Equations of the First Degree in an Unknown Vector

A vector equation involving a single unknown vector, and that not more than once, if at all, in each of its terms, is a vector equation of the first degree in the unknown vector; we suppose the unknown vector to be the only unknown quantity in the equation.

In such an equation, if $\mathbf{v}$ denote the unknown vector, we may have terms of the following types.

(1)	(2)	(3)	(4)
$\mathbf{d}\mathbf{e} \cdot \mathbf{v}$	$s\mathbf{v}$	$\mathbf{g} \times \mathbf{v}$	$\mathbf{h}$.

Terms of types (2) and (3) can be expressed as the sum of three terms of type (1). For, by equation (2′), Art. 20:

$$\mathbf{v} = \mathbf{a}^1\mathbf{a}_1 \cdot \mathbf{v} + \mathbf{a}^2\mathbf{a}_2 \cdot \mathbf{v} + \mathbf{a}^3\mathbf{a}_3 \cdot \mathbf{v},$$

where a_1, a_2, a_3 and a^1, a^2, a^3 are the vectors of reciprocal base-systems; hence:

$$s\mathrm{v} = s\mathrm{a}^1\mathrm{a}_1 \cdot \mathrm{v} + s\mathrm{a}^2\mathrm{a}_2 \cdot \mathrm{v} + s\mathrm{a}^3\mathrm{a}_3 \cdot \mathrm{v};$$
$$\mathrm{g} \times \mathrm{v} = \mathrm{g} \times \mathrm{a}^1\mathrm{a}_1 \cdot \mathrm{v} + \mathrm{g} \times \mathrm{a}^2\mathrm{a}_2 \cdot \mathrm{v} + \mathrm{g} \times \mathrm{a}^3\mathrm{a}_3 \cdot \mathrm{v};$$

the terms on the right of these equations are each of type (1). Suppose, then, that all terms of types (2) and (3) in the vector equation of the first degree in the unknown vector have been thus reduced to terms of type (1). For the typical term of type (1) we can write:

$$\mathrm{de} \cdot \mathrm{v} = \mathrm{d}e_1\,\mathrm{a}_1 \cdot \mathrm{v} + \mathrm{d}e_2\mathrm{a}_2 \cdot \mathrm{v} + \mathrm{d}e_3\,\mathrm{a}_3 \cdot \mathrm{v},$$

where $e_1\mathrm{a}_1$, $e_2\mathrm{a}_2$, $e_3\mathrm{a}_3$ are the components of e on the a_1, a_2, a_3-base-system. Suppose such a reduction to have been made for each term of type (1), and let l_1, l_2, l_3 denote the sum of the cofactors of $\mathrm{a}_1 \cdot \mathrm{v}$, $\mathrm{a}_2 \cdot \mathrm{v}$, $\mathrm{a}_3 \cdot \mathrm{v}$, respectively, in the resulting expression. The equation under consideration can then be put in the form:

$$\mathrm{l}_1\mathrm{a}_1 \cdot \mathrm{v} + \mathrm{l}_2\mathrm{a}_2 \cdot \mathrm{v} + \mathrm{l}_3\mathrm{a}_3 \cdot \mathrm{v} = \mathrm{q},$$

where q includes all terms of type (4). Taking in turn the scalar products of each member of this equation by l^1, l^2, l^3, where the system $(\mathrm{l}^1, \mathrm{l}^2, \mathrm{l}^3)$ is reciprocal to the system $(\mathrm{l}_1, \mathrm{l}_2, \mathrm{l}_3)$, we obtain:

$$\mathrm{a}_1 \cdot \mathrm{v} = \mathrm{q} \cdot \mathrm{l}^1, \qquad \mathrm{a}_2 \cdot \mathrm{v} = \mathrm{q} \cdot \mathrm{l}^2, \qquad \mathrm{a}_3 \cdot \mathrm{v} = \mathrm{q} \cdot \mathrm{l}^3.$$

But, in virtue of equation (2′), Art. 20, we can write:

$$\mathrm{v} = \mathrm{a}_1 \cdot \mathrm{v}\mathrm{a}^1 + \mathrm{a}_2 \cdot \mathrm{v}\mathrm{a}^2 + \mathrm{a}_3 \cdot \mathrm{v}\mathrm{a}^3.$$

Hence, the required solution is given by the equation:

$$\mathrm{v} = \mathrm{q} \cdot \mathrm{l}^1\mathrm{a}^1 + \mathrm{q} \cdot \mathrm{l}^2\mathrm{a}^2 + \mathrm{q} \cdot \mathrm{l}^3\mathrm{a}^3.$$

It has here been tacitly assumed that l_1, l_2, l_3 are non-coplanar vectors. Should this not be the case, these vectors will have no reciprocal system and the method will fail; in this case it would be necessary to resort to special methods.

EXERCISES ON CHAPTER I

§§1–5

1. If the sum and the difference of two line-vectors are given, show graphically how the line-vectors themselves can be found.

2. Verify graphically that the associative and commutative laws are valid in the addition of three coplanar line-vectors.

3. Three line-vectors of lengths a, a, $\sqrt{a}$, respectively, are mutually perpendicular. Determine the magnitude of their resultant and the angle between its direction and that of each of the line-vectors.

4. Under what conditions will a system of vectors of magnitudes 7, 24, and 25, respectively, have a vanishing resultant?

5. If A, B, C, D, E, F represent points at the vertices of a regular hexagon with sides of unit length, find the magnitude and direction of the resultant of the system of forces represented by the line-vectors $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{AD}$, $\overrightarrow{AE}$, $\overrightarrow{AF}$.

6. The point O is any point in the plane of a triangle ABC, and the points D, E, F are middle points of its sides. Show that the system of forces represented by the line-vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$ is equivalent to the system represented by the line-vectors $\overrightarrow{OD}$, $\overrightarrow{OE}$, $\overrightarrow{OF}$.

§§6–10

7. Find the direction and magnitude of a vector whose component measure=numbers on an $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-base-system are 3, 5, and 7, respectively.

8. Show that the vectors

$$\mathfrak{a} = \mathbf{i}S\cos\phi + \mathbf{j}S\sin\phi,$$
$$\mathfrak{b} = \mathbf{i}P\cos\phi + \mathbf{j}P\sin\phi,$$

are parallel.

9. Write out in the form

$$\mathfrak{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

the expressions for three different unit vectors.

10. Write out in the form

$$\mathfrak{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

the expressions for three different pairs of mutually perpendicular vectors.

11. Show that the bisectors of the angles of a triangle meet in a point.

12. Show that a line from a vertex of a parallelogram to the middle point of a non-adjacent side trisects a diagonal.

13. Show that among any four vectors there must exist at least one relation expressible in the form of a vector equation with scalar coefficients. Discuss the cases for which there are, respectively, one, two, and three such relations.

14. Find the vector equation of the line which passes through a given point and which intersects any two given lines.

15. If A, B, C are three weighted points in a plane, show that any point in their plane can be made their centroid, provided the weights of the points can be chosen positive or negative at will.

§§11–15

16. In the orthogonal transformation from an $\mathbf{i}, \mathbf{j}, \mathbf{k}$- to an $\mathbf{i}', \mathbf{j}', \mathbf{k}'$-system of axes expressed by the first set of equations (3), Art. 13, the determinant of the transformation is as follows:

$$\begin{vmatrix} \mathbf{i}\cdot\mathbf{i}' & \mathbf{j}\cdot\mathbf{i}' & \mathbf{k}\cdot\mathbf{i}' \\ \mathbf{i}\cdot\mathbf{j}' & \mathbf{j}\cdot\mathbf{j}' & \mathbf{k}\cdot\mathbf{j}' \\ \mathbf{i}\cdot\mathbf{k}' & \mathbf{j}\cdot\mathbf{k}' & \mathbf{k}\cdot\mathbf{k}' \end{vmatrix}$$

Show that this determinant has the value unity; show also that:

$$\mathbf{k}'\cdot\mathbf{i} = \begin{vmatrix} \mathbf{i}'\cdot\mathbf{j} & \mathbf{i}'\cdot\mathbf{k} \\ \mathbf{j}'\cdot\mathbf{j} & \mathbf{j}'\cdot\mathbf{k} \end{vmatrix}, \qquad \mathbf{k}'\cdot\mathbf{j} = \begin{vmatrix} \mathbf{i}'\cdot\mathbf{k} & \mathbf{i}'\cdot\mathbf{i} \\ \mathbf{j}'\cdot\mathbf{k} & \mathbf{j}'\cdot\mathbf{i} \end{vmatrix}, \qquad \mathbf{k}'\cdot\mathbf{k} = \begin{vmatrix} \mathbf{i}'\cdot\mathbf{i} & \mathbf{i}'\cdot\mathbf{j} \\ \mathbf{j}'\cdot\mathbf{i} & \mathbf{j}'\cdot\mathbf{j} \end{vmatrix},$$

with similar relations for $\mathbf{i}'\cdot\mathbf{i}$, $\mathbf{i}'\cdot\mathbf{j}$, $\mathbf{i}'\cdot\mathbf{k}$ and for $\mathbf{j}'\cdot\mathbf{i}$, $\mathbf{j}'\cdot\mathbf{j}$, $\mathbf{j}'\cdot\mathbf{k}$; finally, show that the vector:

$$\mathbf{i}'\cdot\mathbf{k}\mathbf{j} - \mathbf{i}'\cdot\mathbf{j}\mathbf{k} = \mathbf{j}'\cdot\mathbf{i}\mathbf{k}' - \mathbf{k}'\cdot\mathbf{i}\mathbf{j}'$$

is collinear with the line of intersection of the $\mathbf{j}, \mathbf{k}$-plane and the $\mathbf{j}', \mathbf{k}'$-plane, and that its magnitude is equal to the sine of the angle between these planes.

17. If O designate a fixed point and AB a given line, and if a point Q is taken in the line OP through a point P in the line AB such that

$$\mathbf{r}\cdot\mathbf{q} = k^2,$$

where $\mathbf{r}$, $\mathbf{q}$ denote the position vectors of P, Q with respect to O and k is a constant, find the locus of the point Q.

18. Show that a system of forces represented in magnitude, direction, and position, by the sides of a plane polygon, considered as line-vectors, all directed clockwise (or counter-clockwise) is equivalent to a couple whose moment is equal in magnitude to twice the area of the polygon.

19. Find the point of intersection of a given line and a given plane.

20. Find the equation of the line of intersection of two given planes.

21. Show that:

$$[\mathbf{a}\times\mathbf{b}\ \ \mathbf{b}\times\mathbf{c}\ \ \mathbf{c}\times\mathbf{a}] = [\mathbf{abc}]^2.$$

22. The position vectors of the vertices A, B, C, D of a tetrahedron are given with respect to an $\mathbf{i}, \mathbf{j}, \mathbf{k}$-base-system as follows:

$$\begin{aligned} \mathbf{a} &= x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}, \\ \mathbf{b} &= x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}, \\ \mathbf{c} &= x_3\mathbf{i} + y_3\mathbf{j} + z_3\mathbf{k}, \\ \mathbf{d} &= x_4\mathbf{i} + y_4\mathbf{j} + z_4\mathbf{k}, \end{aligned}$$

where the x, y, z-measure-numbers are rectangular Cartesian co-ordinates of the vertices. If V denote the volume of the tetrahedron, show vectorially that:

$$V = \frac{1}{6}\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix}$$

23. Show that:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0.$$

24. Show that:

$$[\mathbf{pqr}]\,[\mathbf{abc}] = \begin{vmatrix} \mathbf{p}\cdot\mathbf{a} & \mathbf{p}\cdot\mathbf{b} & \mathbf{p}\cdot\mathbf{c} \\ \mathbf{q}\cdot\mathbf{a} & \mathbf{q}\cdot\mathbf{b} & \mathbf{q} \\ \mathbf{r}\cdot\mathbf{a} & \mathbf{r}\cdot\mathbf{b} & \mathbf{r} \end{vmatrix}$$

25. Show that:

$$[\mathbf{pqr}]\,(\mathbf{a} \times \mathbf{b}) \qquad \begin{vmatrix} \mathbf{p}\cdot\mathbf{a} & \mathbf{p}\cdot\mathbf{b} & \mathbf{p} \\ \mathbf{q}\cdot\mathbf{a} & \mathbf{q}\cdot\mathbf{b} & \mathbf{q} \\ \mathbf{r}\cdot\mathbf{a} & \mathbf{r}\cdot\mathbf{b} & \mathbf{r} \end{vmatrix}$$

26. Deduce the fundamental formulas of spherical trigonometry from the relations:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= \mathbf{a}\cdot\mathbf{c}\,\mathbf{b}\cdot\mathbf{d} - \mathbf{a}\cdot\mathbf{d}\,\mathbf{b}\cdot\mathbf{c}, \\ (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= [\mathbf{acd}]\,\mathbf{b} - [\mathbf{bcd}]\,\mathbf{a} \\ &= [\mathbf{abd}]\,\mathbf{c} - [\mathbf{abc}]\,\mathbf{d}. \end{aligned}$$

Hint: Assume $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ to be unit vectors represented by line-vector radii of a sphere of unit radius.

27. Let $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, $\mathbf{r}_4$ denote the position vectors of any four non-coplanar points P_1, P_2, P_3, P_4, with respect to an arbitrary origin O. Define three non-coplanar vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, as follows:

$$\mathbf{a}_1 = \mathbf{r}_2 - \mathbf{r}_1, \qquad \mathbf{a}_2 = \mathbf{r}_3 - \mathbf{r}_1, \qquad \mathbf{a}_3 = \mathbf{r}_4 - \mathbf{r}_1,$$

and let $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ denote the vectors of the system reciprocal to the $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$-system. Show that the distance of the point P from the plane determined by the points P_1, P_2, P_3 is given by the expression

$$\frac{\mathbf{a}^3 \cdot \mathbf{a}_1}{\sqrt{\mathbf{a}^3 \cdot \mathbf{a}^3}},$$

also show that the shortest distance between the two lines P_1P_2 and P_3P_4 is given by the expression

$$\frac{\mathbf{a}^3 \cdot \mathbf{a}_1}{\sqrt{(\mathbf{a}^2 + \mathbf{a}^3) \cdot (\mathbf{a}^2 + \mathbf{a}^3)}}.$$

CHAPTER II

THE ELEMENTS OF VECTOR CALCULUS

§23

Vector Differentiation

The basic principles underlying the operations of differentiation and integration in vector calculus are similar to the corresponding principles of diffentiation and integration in ordinary calculus, a subject with which the reader will be assumed already familiar.

We consider a vector $\mathbf{v}$ which is supposed related to a scalar parameter s in such manner that as s varies continuously so also does $\mathbf{v}$. The dependency of $\mathbf{v}$ upon s is indicated by writing:

(1) $$\mathbf{v} = \mathbf{v}(s).$$

If $\Delta\mathbf{v}$ *denote the increment in the vector function* $\mathbf{v}$ *due to an increment* Δs *of the parameter* s*, then the vector* $d\mathbf{v}/ds$ *defined as follows:*

(2) $$\frac{d\mathbf{v}}{ds} = \underset{\Delta s \to 0}{Lt} \frac{\Delta \mathbf{v}}{\Delta s} = \underset{\Delta s \to 0}{Lt} \frac{\mathbf{v}(s + \Delta s) - \mathbf{v}(s)}{\Delta s},$$

is called the Derivative of the vector $\mathbf{v}$ *with respect to the parameter* s.

The derivative of the product of a constant and a vector function of a variable parameter is by this definition equal to the product of the constant and the derivative of the vector.

The vector $\mathbf{v}$ can be expressed in terms of its components on a fixed $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$-base-system by writing:

(3) $$\mathbf{v} = x\mathfrak{a} + y\mathfrak{b} + z\mathfrak{c},$$

where the measure-numbers x, y, z of the components are continuous scalar functions of the parameter s. By successive differentiation, assuming x, y, z differentiable to the n'th order, we find:

$$\frac{d\mathbf{v}}{ds} = \frac{dx}{ds}\,\mathfrak{a} + \frac{dy}{ds}\,\mathfrak{b} + \frac{dz}{ds}\,\mathfrak{c},$$

$$\frac{d^2\mathbf{v}}{ds^2} = \frac{d^2x}{ds^2}\,\mathfrak{a} + \frac{d^2y}{ds^2}\,\mathfrak{b} + \frac{d^2z}{ds^2}\,\mathfrak{c},$$

(4)

$$\frac{d^n\mathbf{v}}{ds^n} = \frac{d^nx}{ds^n}\,\mathfrak{a} + \frac{d^ny}{ds^n}\,\mathfrak{b} + \frac{d^nz}{ds^n}\,\mathfrak{c}.$$

If u is a continuous scalar function of a parameter s and $\mathbf{v}$ a continuous vector function of the same parameter, then:

(5) $$\frac{\text{[illegible]}}{ds} = \underset{\Delta s \to 0}{Lt} \frac{\text{[illegible]}}{\Delta s}$$

and hence:

(6) $$\frac{d(u\mathbf{v})}{ds} = u\frac{d\mathbf{v}}{ds} + \frac{du}{ds}\,\mathbf{v}.$$

This formula is precisely analogous to that in ordinary calculus for the derivative of the product of two scalar quantities each of which is a function of a given parameter.

Consider now a line-vector $\underline{\mathbf{r}}$, drawn from a fixed point O to represent a vector $\mathbf{r}$ which is a continuous function of a single parameter s. As s varies continuously, the terminal point of $\mathbf{r}$ will trace out a curve C in space, shown in Fig. 21. When the parameter s increases by the amount Δs, the vector $\mathbf{r}$ represented by $\underline{\mathbf{r}}$ will change by a corresponding amount, represented in the figure by $\Delta\underline{\mathbf{r}}$, a vector chord of the curve C from the terminal point of $\underline{\mathbf{r}}(s)$ to the terminal point of $\underline{\mathbf{r}}(s + \Delta s)$. As Δs approaches zero, $\Delta\underline{\mathbf{r}}$ will tend toward coincidence in direction with the tangent to the curve C at the extremity of $\underline{\mathbf{r}}(s)$ and, therefore, the vector derivative $d\mathbf{r}/ds$ will be a vector whose direction is that of a tangent to the curve in the direction of s increasing.

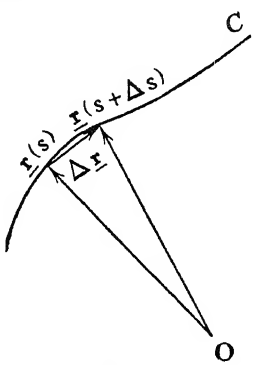


Fig. 21.

As a special example we take the case in which the curve C is represented by the equation:

(7) $$\mathbf{r} = \mathfrak{a}\cos s + \mathfrak{b}\sin s,$$

where $\mathfrak{a}$ and $\mathfrak{b}$ are constant vectors and s a parametric angle. By differentiation with respect to s we get:

$$\frac{d\mathfrak{r}}{ds} = -\mathfrak{a} \sin s + \mathfrak{b} \cos s.$$

It is evident that:

$$\begin{aligned} \mathfrak{r} &= \mathfrak{a} \\ \frac{d\mathfrak{r}}{ds} &= \mathfrak{b} \end{aligned} \quad \text{for } s = 0, \qquad \frac{d\mathfrak{r}}{ds} = -\mathfrak{a} \quad \text{for } s = \frac{\pi}{2}.$$

These equations show that the tangent to the curve at the extremity of $\mathfrak{a}$ is parallel to $\mathfrak{b}$, and vice versa. The curve is an ellipse with $\mathfrak{a}$ and $\mathfrak{b}$, as vectorial conjugate radii.

If the parametric angle s be taken proportional to the time parameter t, we can write $s = \omega t$ where ω represents a scalar factor of proportionality, and the terminal point of the vector $\mathfrak{r}$ will move in the ellipse as orbit with the vector-velocity:

$$\text{(9)} \qquad \frac{d\mathfrak{r}}{dt} = \omega(-\mathfrak{a} \sin \omega t + \mathfrak{b} \cos \omega t);$$

and, with the vector-acceleration:

$$\text{(10)} \qquad \frac{d^2\mathfrak{r}}{dt^2} = -\omega^2(\mathfrak{a} \cos \omega t + \mathfrak{b} \sin \omega t) = -\omega^2\mathfrak{r}.$$

We consider next the differentiation of the scalar product of two vectors $\mathfrak{u}$, $\mathfrak{v}$, each of which is a continuous function of a single scalar parameter s. We have:

$$\frac{d(\mathfrak{u} \cdot \mathfrak{v})}{ds} = \underset{\Delta s \to 0}{\mathrm{Lt}} \frac{(\mathfrak{u} + \Delta\mathfrak{u}) \cdot (\mathfrak{v} + \Delta\mathfrak{v}) - \mathfrak{u} \cdot \mathfrak{v}}{\Delta s},$$

and hence:

$$\text{(11)} \qquad \frac{d(\mathfrak{u} \cdot \mathfrak{v})}{ds} = \mathfrak{u} \cdot \frac{d\mathfrak{v}}{ds} + \frac{d\mathfrak{u}}{ds} \cdot \mathfrak{v}.$$

In like manner it can be shown that, if $\mathfrak{u}$, $\mathfrak{v}$, $\mathfrak{w}$ are continuous functions of a single scalar parameter s, we shall have:

$$\text{(12)} \qquad \frac{d}{ds}(\mathfrak{u} \cdot \mathfrak{v} \times \mathfrak{w}) = \mathfrak{u} \cdot \mathfrak{v} \times \frac{d\mathfrak{w}}{ds} + \mathfrak{u} \cdot \frac{d\mathfrak{v}}{ds} \times \mathfrak{w} + \frac{d\mathfrak{u}}{ds} \cdot \mathfrak{v} \times \mathfrak{w},$$

$$\text{(13)} \qquad \frac{d}{ds}[\mathfrak{u} \times (\mathfrak{v} \times \mathfrak{w})] = \mathfrak{u} \times \left(\mathfrak{v} \times \frac{d\mathfrak{w}}{ds}\right) + \mathfrak{u} \times \left(\frac{d\mathfrak{v}}{ds} \times \mathfrak{w}\right) + \frac{d\mathfrak{u}}{ds} \times (\mathfrak{v} \times \mathfrak{w}),$$

for the derivatives of their scalar triple product and their vector triple product respectively.

A continuous vector function of more than one independent scalar parameter may be differentiated with respect to any one of them on the assumption that all the others are held constant. The resulting derivative will be the partial derivative of the vector function with respect to the parameter which is supposed to vary.

A simple example will suffice to exemplify the process. Suppose a vector $\mathbf{v}$ to be expressed in terms of any three constant non-coplanar vectors $\mathfrak{a}$, $\mathbf{b}$, $\mathbf{c}$ as follows:

$$\mathbf{v} = x\mathfrak{a} + y\mathbf{b} + z\mathbf{c},$$

where x, y, z are independent variable parameters. Partial differentiation with respect to x, y, z in turn gives:

$$\frac{\partial \mathbf{v}}{\partial x} = \mathfrak{a}, \qquad \frac{\partial \mathbf{v}}{\partial y} = \mathbf{b}, \qquad \frac{\partial \mathbf{v}}{\partial z} = \mathbf{c};$$

and further partial differentiation gives:

$$\frac{\partial^2 \mathbf{v}}{\partial x^2} = 0, \qquad \frac{\partial^2 \mathbf{v}}{\partial y^2} = 0, \qquad \frac{\partial^2 \mathbf{v}}{\partial z^2} = 0.$$

It is often convenient in vector calculus as in ordinary calculus to operate with differentials rather than with derivatives. The total differential of the continuous vector function:

$$\mathbf{v} = x\mathfrak{a} + y\mathbf{b} + z\mathbf{c},$$

can, for example, be expressed in the form:

$$d\mathbf{v} = dx\mathfrak{a} + dy\mathbf{b} + dz\mathbf{c}, \tag{14}$$

where dx, dy, dz are independent differentials of the variable scalar parameters x, y, z.

§24
Vector Integration

The process of vector integration is the inverse of vector differentiation.

Consider a vector $\mathbf{v}(s)$ which depends on a single scalar parameter s, and suppose it to be a finite, single-valued, continuous function of s.[1] By an integral, say $\mathfrak{q}(s)$, of the vector $\mathbf{v}$ is meant

[1] All functions here dealt with are supposed subject to these conditions; cases in which they are not satisfied must receive special treatment just as in corresponding cases in ordinary calculus, with which the reader is supposed familiar.

a vector function of s which when differentiated with respect to s gives the vector $\mathbf{v}$. Since the differential of a constant vector is zero, the result of adding any constant vector $\mathbf{c}$ to the integral $\mathbf{q}(s)$ will also be an integral of the vector $\mathbf{v}$ or, in other words, the integral is only determinate to a constant vector. Hence, the *indefinite* integral of $\mathbf{v}$ is expressed by writing:

$$\int \mathbf{v}(s)\,ds = \mathbf{q}(s) + \mathbf{c}. \tag{1}$$

If the parameter s vary continuously from a definite value a to a definite value b, then:

$$\int_a^b \mathbf{v}(s)\,ds = \mathbf{q}(b) - \mathbf{q}(a) \tag{2}$$

expresses the *definite* integral of the vector $\mathbf{v}$ between the limits a and b; it represents the vector sum of the vector differential increments $\mathbf{v}(s)ds$ ($= d\mathbf{q}(s)$) as s varies from a to b.

Suppose $\mathbf{v}(s)$, $\mathbf{q}(s)$, and $\mathbf{c}$ to be expressed in terms of any three constant non-coplanar vectors ($\mathbf{l}$, $\mathbf{m}$, $\mathbf{n}$) as follows:

$$\begin{aligned} \mathbf{v}(s) &= v_l\mathbf{l} + v_m\mathbf{m} + v_n\mathbf{n}, \\ \mathbf{q}(s) &= q_l\mathbf{l} + q_m\mathbf{m} + q_n\mathbf{n}, \\ \mathbf{c} &= c_l\mathbf{l} + c_m\mathbf{m} + c_n\mathbf{n}, \end{aligned}$$

where, of course, the coefficients of $\mathbf{l}$, $\mathbf{m}$, $\mathbf{n}$ in the equations for $\mathbf{v}$ and $\mathbf{q}$ are finite, single-valued, continuous functions of the variable parameter s. Then, instead of (1) and (2) we can write:

$$(3)\quad \begin{aligned} \int v_l ds &= q_l(s) + c_l, \\ \int v_m ds &= q_m(s) + c_m, \\ \int v_n ds &= q_n(s) + c_n; \end{aligned} \qquad (4)\quad \begin{aligned} \int_a^b v_l ds &= q_l(b) - q_l(a), \\ \int_a^b v_m ds &= q_m(b) - q_m(a), \\ \int_a^b v_n ds &= q_n(b) - q_n(a). \end{aligned}$$

The process of vector integration can thus be reduced to scalar integration.

Now suppose the vector $\mathbf{v}(s)$ to be affected by a scalar coefficient $f(s)$, subject to the same restrictions as $\mathbf{v}(s)$, and that the value of the integral

$$\int f(s)\,\mathbf{v}(s)\,ds$$

is required. The method of integration by parts can be applied in two ways as follows: If $\mathbf{q}(s)$ is an integral of $\mathbf{v}(s)$ and $F(s)$ an integral of $f(s)$, then:

$$\int f(s)\,\mathbf{v}(s)\,ds = f(s)\,\mathbf{q}(s) - \int f'(s)\,\mathbf{q}(s)\,ds, \tag{5}$$

$$\int f(s)\,\mathbf{v}(s)\,ds = F(s)\,\mathbf{v}(s) - \int F(s)\,\mathbf{v}'(s)\,ds, \tag{6}$$

as is at once evident upon differentiating the right-hand members of these equations with respect to s.

The validity of the following equations is evident upon inspection:

$$\int \mathbf{a}\cdot\mathbf{v}ds = \mathbf{a}\cdot\int\mathbf{v}ds, \tag{7}$$

$$\int \mathbf{a}\times\mathbf{v}ds = \mathbf{a}\times\int\mathbf{v}ds, \tag{8}$$

where $\mathbf{a}$ is a vector which does not depend upon the parameter s.

§25
Line, Surface, and Volume Integrals Involving Vectors

Let $\mathbf{v}$ denote a vector function of the co-ordinates of any point P and $\mathbf{r}$ the position-vector of the point. Let $d\mathbf{r}$ denote a differential increment of the position-vector $\mathbf{r}$. If A and B denote any two points, then the integral

$$\int_A^B \mathbf{v}\cdot d\mathbf{r},$$

taken along some path connecting A and B is a scalar line integral involving $\mathbf{v}$. If s denote the numerical measure of the distance along the path from A (or some other fiducial point) to the point P, and if the position-vector $\mathbf{r}$ be considered as a function of s as a parameter, then $d\mathbf{r}/ds$ will be a unit vector ($\mathbf{t}$) whose direction is that of a tangent to the path at P in the direction of increasing s. We can then write:

$$\int_A^B \mathbf{v}\cdot d\mathbf{r} = \int_A^B \mathbf{t}\cdot\mathbf{v}ds. \tag{1}$$

The line integral in question is therefore nothing more than an ordinary integral along the path from A to B of the component of $\mathbf{v}$ tangential to the path.

A line integral will in general depend upon the path as well as the initial and terminal point of the path. We shall consider later the conditions under which the value of a line integral is independent of the path when its initial and terminal points are given.

Consider now a surface S, closed or un-closed, with each point of which is associated a value of a vector function $\mathbf{v}$ of the co-ordinates of the point, and call one of its sides positive and the other negative. Let $\mathbf{n}$ denote a unit vector whose direction is that of a normal to the positive side (determined by convention) of the surface. Then the scalar value at this element of the normal component of $\mathbf{v}$ in the direction of $\mathbf{n}$ will be $\mathbf{n} \cdot \mathbf{v}$, and the scalar integral

$$\int_S \mathbf{n} \cdot \mathbf{v} d\sigma,$$

where $d\sigma$ denotes the magnitude of an element of area of S, taken over the surface S represents the surface integral of the normal component of $\mathbf{v}$, a quantity which is sometimes called the Flux of the Vector $\mathbf{v}$ through the surface S from its negative to its positive side.

As an example of a line integral which itself defines a vector, we take the case of the integral of a scalar function of position along a space curve. If f denote the function and $\mathbf{r}$ the position-vector of a point on the curve, then the integral

$$\int_A^B f d\mathbf{r}$$

is a vector which represents the vector sum of $f d\mathbf{r}$ along the curve between two points A and B.

Similarly, if f denote a scalar point function and $\mathbf{n}$ a unit vector directed normally to the positive side of a surface S for which $d\sigma$ is the magnitude of the typical element of area, the integral

$$\int f \mathbf{n} d\sigma$$

taken over the surface S represents a vector.

As a final example, consider a region of space V in the field of a vector point function $\mathbf{v}$, and suppose V to be subdivided into differential elements of magnitude $d\tau$. The integral

$$\int \mathbf{v} d\tau$$

taken throughout V is a volume integral which itself represents a vector.

In the following pages many examples of line, surface and volume integrals will be found, and in Chapter IV several transformation theorems relating to such integrals will be derived.

§26[1)]

Rotating Axes

We consider two systems of axes with a common origin O designated respectively as the A-system and the B-system, and suppose the two systems to be in relative motion but with the common origin O fixed.

Let $\mathbf{q}$ denote the position-vector of a point Q fixed in the B-system, and t a time parameter, then the vector $d\mathbf{q}/dt$ will represent the velocity of Q with respect to the A-system, if $d\mathbf{q}$ denote the differential increment of $\mathbf{q}$ corresponding to the differential increment dt in the time parameter for an observer A fixed in the A-system. For an observer B fixed in the B-system the velocity of Q will of course be zero, since for him Q is a fixed point.

The B-system may evidently be considered as fixed in a rotating rigid body whose instantaneous axis passes through the fixed point O. Hence, by equation (5), Art. 15:

$$\frac{d\mathbf{q}}{dt} = \boldsymbol{\omega} \times \mathbf{q}, \tag{1}$$

where the vector $\boldsymbol{\omega}$ specifies the angular velocity of the B-system with respect to the A-system.

Let $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ denote the base-vectors of a unit orthogonal right-handed base-system fixed in the B-system. The angular velocity-vector $\boldsymbol{\omega}$ can be expressed in terms of its components on this base-system by writing:

$$\boldsymbol{\omega} = \omega_1\mathbf{b}_1 + \omega_2\mathbf{b}_2 + \omega_3\mathbf{b}_3, \tag{2}$$

[1)] The present article and the remaining articles of the present chapter are concerned with various geometrical and physical applications of vector calculus. As in preceding applications of vector methods given in the present book, the symbols representing geometrical or physical quantities, whose relationships are objects of investigation, generally denote pure scalars or pure vectors. It is, however, often more convenient to consider the symbols used to represent geometrical or physical quantities as carrying the physical dimensions of these quantities. If this procedure is followed, care must be taken to ensure that all equations introduced shall balance dimensionally.

where the ω-coefficients are the measure-numbers of the components. From equations (1) and (2), by taking $\mathbf{q}$ as $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ in turn, we find:

$$\frac{d\mathbf{b}_1}{dt} = \omega_3\mathbf{b}_2 - \omega_2\mathbf{b}_3,$$

$$\frac{d\mathbf{b}_2}{dt} = \omega_1\mathbf{b}_3 - \omega_3\mathbf{b}_1,$$

$$\frac{d\mathbf{b}_3}{dt} = \omega_2\mathbf{b}_1 - \omega_1\mathbf{b}_2.$$

Upon multiplication of these equations by $\mathbf{b}_2\cdot$, $\mathbf{b}_3\cdot$, $\mathbf{b}_1\cdot$, respectively, we get for the measure-numbers of the vector $\boldsymbol{\omega}$:

$$\omega_1 = \mathbf{b}_3\cdot\frac{d\mathbf{b}_2}{dt},$$

$$\omega_2 = \mathbf{b}_1\cdot\frac{d\mathbf{b}_3}{dt},$$

$$\omega_3 = \mathbf{b}_2\cdot\frac{d\mathbf{b}_1}{dt}.$$

Hence, for the vector $\boldsymbol{\omega}$ itself we can write:

(3) $$\boldsymbol{\omega} = \mathbf{b}_3\cdot\frac{d\mathbf{b}_2}{dt}\,\mathbf{b}_1 + \mathbf{b}_1\cdot\frac{d\mathbf{b}_3}{dt}\,\mathbf{b}_2 + \mathbf{b}_2\cdot\frac{d\mathbf{b}_1}{dt}\,\mathbf{b}_3,$$

an equation which will be of use later.

§27
Relative Motion Due to Rotation

Now suppose a point P with position-vector $\mathbf{r}$ to be in motion with respect to both the A and the B-system. The velocity-vector $(d\mathbf{r}/dt)$ for P with respect to the A-system will, of course, be different in general from its velocity-vector $(\delta\mathbf{r}/\delta t)$ with respect to the B-system. The relation between these two velocity-vectors is expressed by the equation:[1]

(1) $$\frac{d\mathbf{r}}{dt} = \frac{\delta\mathbf{r}}{\delta t} + \boldsymbol{\omega}\times\mathbf{r};$$

for the velocity of the point P with respect to the A-system must evidently be equal to its velocity with respect to the B-system increased by the velocity which it would have if it be supposed fixed at the instant under consideration with respect to the B-system, and this by equation (1), Art. 24, is specified by the vector $\boldsymbol{\omega}\times\mathbf{r}$.

[1] It is assumed the units of measure for observers A and B are the same.

From a purely kinematical point of view the two systems are on a basis of parity in the sense that the symbols of differentiation d and δ can be interchanged, provided at the same time the angular velocity-vector $(-\boldsymbol{\omega})$ be substituted for the angular velocity-vector $\boldsymbol{\omega}$.

If the vector $\mathbf{r}$ specify in magnitude and direction any vector quantity $\mathbf{R}$, whatever its nature, then, by equation (1):

$$\frac{d\mathbf{R}}{dt} = \frac{\delta\mathbf{R}}{\delta t} + \boldsymbol{\omega} \times \mathbf{R}. \tag{2}$$

This equation gives the relation of the estimates of the changes per unit time in $\mathbf{R}$ which would be made by the observers A and B.

For example, if $\mathbf{R} = \boldsymbol{\omega}$, it follows from the last equation, since $\boldsymbol{\omega} \times \boldsymbol{\omega} = 0$, that:

$$\frac{d\boldsymbol{\omega}}{dt} = \frac{\delta\boldsymbol{\omega}}{\delta t}. \tag{3}$$

This equation implies that the estimates of the angular acceleration of the B-system with respect to the A-system as made by A and B would be the same.

As another example, if we take $\mathbf{R} = d\mathbf{r}/dt$, we obtain by equation (2):

$$\frac{d}{dt}\frac{d\mathbf{r}}{dt} = \frac{\delta}{\delta t}\frac{d\mathbf{r}}{dt} + \boldsymbol{\omega} \times \frac{d\mathbf{r}}{dt}.$$

Making use of equation (1), this equation can be written in the form:

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{\delta^2\mathbf{r}}{\delta t^2} + 2\boldsymbol{\omega} \times \frac{\delta\mathbf{r}}{\delta t} + \frac{\delta\boldsymbol{\omega}}{\delta t} \times \mathbf{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}). \tag{4}$$

The term on the left and the first term on the right of this equation specify, respectively, the total acceleration of the point P as estimated by observers A and B. The second term on the right specifies a contribution to the total acceleration observed by A which is commonly called the Compound Acceleration of Coriolis; its direction is perpendicular to the directions of the instantaneous axis and of the velocity of the point P as seen by observer B. The sum of the two remaining terms on the right specifies a contribution to the total acceleration of P observed by A which is sometimes called the Acceleration of Moving Space;[1)] for, if the point P is supposed fixed in the B-system, the sum of these two

[1)] This terminology implies that the space of A is considered as fixed and that of B as moving.

terms would represent the total acceleration observed by A, since the first two terms on the right of equation (4) would then vanish; in the particular case of uniform rotation the acceleration of moving space is represented by the single term $\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ which is equivalent to $\boldsymbol{\omega} \cdot \mathbf{r}\boldsymbol{\omega} - \boldsymbol{\omega} \cdot \boldsymbol{\omega}\mathbf{r}$ or $-\omega^2\mathbf{p}$, if $\mathbf{p}$ specify a vector-perpendicular from the instantaneous axis to the point P, whose position-vector is $\mathbf{r}$, and the acceleration of moving space in this case is nothing more than the familiar Centripetal Acceleration of a point rotating uniformly about a fixed axis.

§28

Theory of Curves in Space—Frenet's Formulas

Referring to Fig. 22, let $\underline{\mathbf{r}}$ be a line-vector representing the position-vector of a point P on a space curve with respect to an arbitrary origin O, and s a parameter of the curve representing the distance of P measured along the curve from a fiducial point P_0. The vector $\mathbf{r}$ represented by $\mathbf{r}$ can be considered as dependent upon the parameter s, and it together with its first and second derivatives with respect to s will be considered as differentiable functions of s.

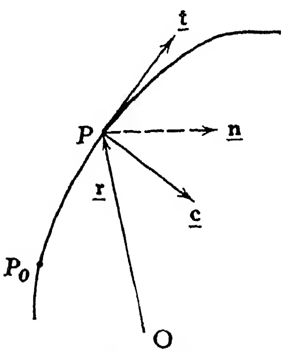


Fig. 22.

Let $\underline{\mathbf{t}}$ represent a unit vector $\mathbf{t}$ at $P(s)$ in the direction of increasing s; then:

(1) $$\mathbf{t} = \frac{d\mathbf{r}}{ds},$$

where $d\mathbf{r}$ represents the increment in $\mathbf{r}$ in passing from the point $P(s)$ to the point $P'(s + ds)$.

We define a unit vector:

(2) $$d\mathbf{t}$$

where $d\mathbf{t}$ denotes the increment in $\mathbf{t}$ in passing from the point $P(s)$ to the point $P'(s + ds)$, and k a positive factor of proportionality; since $\mathbf{t}$ is a unit vector, $d\mathbf{t}$ will be perpendicular to $\mathbf{t}$.

The two vectors $\mathbf{t}$ and $\mathbf{c}$ determine a plane at P called the Osculating Plane of the curve.

Finally, let $\underline{\mathbf{n}}$ represent a unit vector $\mathbf{n}$ at P, perpendicular to the osculating plane in a direction such that:

(3) $$\mathbf{t} \times \mathbf{c}.$$

The three unit vectors **t**, **c**, **n** constitute a unit right-handed orthogonal system associated with the point P, **t** being a vector tangent to the curve, **c** the principal vector normal, and **n** a vector bi-normal.

In passing along the curve from the point $P(s)$ to the neighboring point $P'(s + ds)$, the system at P may be supposed to move so as to bring it into the configuration appropriate to P'. A vector **u** at P fixed in this system will thereby experience in general an infinitesimal translation and rotation, and if ω denote the angular velocity of the rotation and $d\mathbf{u}$ the change in **u** due to the rotation only, then, by equation (1), Art. 26:

$$\frac{d\mathbf{u}}{dt} = \omega \times \mathbf{u},$$

where, by equation (3) of the same article:

$$\omega = \mathbf{n} \cdot \frac{d\mathbf{c}}{dt}\mathbf{t} + \mathbf{t} \cdot \frac{d\mathbf{n}}{dt}\mathbf{c} + \mathbf{c} \cdot \frac{d\mathbf{t}}{dt}\mathbf{n}.$$

Hence, upon introducing the curve parameter s in place of the time parameter t as independent variable, we get:

$$\frac{d\mathbf{u}}{ds} = \Omega \times \mathbf{u}, \tag{4}$$

where:

$$\mathbf{\Omega} = \mathbf{n} \cdot \frac{d\mathbf{c}}{ds}\mathbf{t} + \mathbf{t} \cdot \frac{d\mathbf{n}}{ds}\mathbf{c} + \mathbf{c} \cdot \frac{d\mathbf{t}}{ds}\mathbf{n}. \tag{5}$$

The vector $\mathbf{\Omega}$ is called the Darboux Vector of the curve.

Since $\mathbf{t} \cdot \mathbf{c} = \mathbf{c} \cdot \mathbf{n} = \mathbf{n} \cdot \mathbf{t} = 0$, and since, by equation (2), $d\mathbf{t}$ is parallel to **c**, we have, by differentiation, the following relations:

$$\mathbf{t} \cdot d\mathbf{c} = -\mathbf{c} \cdot d\mathbf{t}, \qquad \mathbf{c} \cdot d\mathbf{n} = -\mathbf{n} \cdot d\mathbf{c}, \qquad \mathbf{n} \cdot d\mathbf{t} = -\mathbf{t} \cdot d\mathbf{n} = 0.$$

The last of these equations shows that $d\mathbf{n}$ must be perpendicular to **t** and, since **n** is a unit vector, it must also be perpendicular to **n**; it therefore follows that $d\mathbf{n}$ must be collinear with **c**. The Darboux vector can now be written as follows:

$$\mathbf{\Omega} = -\mathbf{c} \cdot \frac{d\mathbf{n}}{ds}\mathbf{t} + \mathbf{c} \cdot \frac{d\mathbf{t}}{ds}\mathbf{n}, \tag{6}$$

showing that $\mathbf{\Omega}$ is coplanar with **t** and **n**.

We now define two characteristic quantities associated with the curve at P by the equations:

$$C_1 = \mathbf{c} \cdot \frac{d\mathbf{t}}{ds}; \tag{7}$$

$$C_2 = -\mathbf{c} \cdot \frac{d\mathbf{n}}{ds}. \tag{8}$$

Here, C_1 is equal to the reciprocal of the positive factor of proportionality k in equation (2), and is called the First Curvature or Flexure of the curve at P; C_2 is called the Second Curvature or Torsion of the curve at P.[1)]

C_1 is always positive and, since $\mathbf{c}$ is a unit vector collinear with $d\mathbf{t}$, is equal to $|d\mathbf{t}/ds|$ or, since $\mathbf{t}$ is a unit vector, to the angle turned through by the tangent to the curve per unit distance traversed along it.

C_2 can be either positive or negative according as $d\mathbf{n}$, which is collinear with $\mathbf{c}$, is oppositely or like-directed to $\mathbf{c}$; furthermore, since $\mathbf{c}$ is a unit vector, C_2 is equal to $\mp\, |d\mathbf{n}/ds|$ and, therefore, since $\mathbf{n}$ is a unit vector, is numerically equal to the angle turned through by the osculating plane per unit distance traversed along the curve.

A third characteristic quantity associated with the curve at P is defined by the equation:

$$C = \left|\frac{d\mathbf{c}}{ds}\right|, \tag{9}$$

and, for a reason which will be given presently, is called the Total Curvature.

The Darboux vector by equations (6), (7), and (8) can now be expressed in the simple form:

$$\mathbf{\Omega} = C_2\mathbf{t} + C_1\mathbf{n}. \tag{10}$$

From equations (4) and (10), by taking in turn $\mathbf{t}$, $\mathbf{c}$, $\mathbf{n}$ for the vector $\mathbf{u}$, we get:

$$\begin{aligned} \frac{d\mathbf{t}}{ds} &= \mathbf{\Omega}\times\mathbf{t} = 0\mathbf{t} + C_1\mathbf{c} + 0\mathbf{n}, \\ \frac{d\mathbf{c}}{ds} &= \mathbf{\Omega}\times\mathbf{c} = -C_1\mathbf{t} + 0\mathbf{c} + C_2\mathbf{n}, \\ \frac{d\mathbf{n}}{ds} &= \mathbf{\Omega}\times\mathbf{n} = 0\mathbf{t} - C_2\mathbf{c} + 0\mathbf{n}. \end{aligned} \tag{11}$$

[1)] Strictly, since we are considering the curve parameter s as a pure number, C_1 and C_2 are the numerical magnitudes of the flexure and torsion of the curve at P rather than the torsion and flexure themselves, which have the physical dimensions of an inverse length.

These equations are known as Frenet's Formulas. From them all the geometrical properties of a space curve can be derived.

Upon squaring both members of the second of equations (11) and taking account of equation (9), we get:

(12) $$C^2 = C_1^2 + C_2^2.$$

This equation is known as Lancret's Law, and it is in virtue of this law that the quantity C is called the total curvature of the curve.

The three quantities ρ, ρ_1, ρ_2 defined by the equations

(13) $$\rho = \frac{1}{C}, \qquad \rho_1 = \frac{1}{C_1}, \qquad \rho_2 = \frac{1}{C_2},$$

are the magnitudes of the Radii of Curvature of the curve at P. The point at a distance ρ_1 from P in the direction of $\mathbf{c}$ is called the First Center of curvature, and that at a distance ρ_2 in the direction $(\pm\mathbf{c})$ of $d\mathbf{n}$ is called the Second Center of Curvature of the curve at P.

Expressions for the flèxure and the torsion of a curve in space at any point $P(s)$, whose position-vector is denoted by $\mathbf{r}$, in terms of derivatives of $\mathbf{r}$ with respect to the curve parameter s, can be directly obtained from Frenet's formulas as follows:

Squaring both members of the first of equations (11), we get:

$$C_1^2 = \frac{d\mathbf{t}}{ds}\cdot\frac{d\mathbf{t}}{ds},$$

and, upon taking the square root and using equation (1), we find:

(14)

Multiplying the second of equations (11) by $\mathbf{n}\cdot$, we get:

$$C_2 = \mathbf{n}\cdot\frac{d\mathbf{c}}{ds} = \mathbf{t}\times\mathbf{c}\cdot\frac{d\mathbf{c}}{ds} = \mathbf{t}\cdot\mathbf{c}\times\frac{d\mathbf{c}}{ds};$$

but, taking account of equations (1), (2), and (7):

$$\mathbf{t}\cdot\mathbf{c}\times\frac{d\mathbf{c}}{ds} = \frac{d\mathbf{r}}{ds}\cdot\frac{1}{C_1}\frac{d^2\mathbf{r}}{ds^2}\times\frac{d}{ds}\left(\frac{1}{C_1}\frac{d^2\mathbf{r}}{ds^2}\right)$$
$$= \frac{1}{C_1{}^2}\frac{d\mathbf{r}}{ds}\cdot\frac{d^2\mathbf{r}}{ds^2}\times\frac{d^3\mathbf{r}}{ds^3};$$

hence, taking account of equation (14):

(15) $$C_2 \quad \frac{\frac{d\mathbf{r}}{ds}\cdot\frac{d^2\mathbf{r}}{ds^2}\times\frac{d^3\mathbf{r}}{ds^3}}{\frac{d^2\mathbf{r}}{ds^2}\cdot\frac{d^2\mathbf{r}}{ds^2}}$$

If the position-vector $\mathbf{r}$ of the point P be referred to a fixed $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-base-system with origin at O, then:

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k},$$

where x, y, z are Cartesian co-ordinates of the point P. Using this expression for $\mathbf{r}$, we get directly from equations (14) and (15) their Cartesian equivalents:

(16) $$C_1 = \sqrt{\left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2 + \left(\frac{d^2z}{ds^2}\right)^2},$$

and:

(17) $$\begin{array}{ccc} \dfrac{dx}{ds} & \dfrac{dy}{ds} & \dfrac{dz}{ds} \\ \dfrac{d^2x}{ds^2} & \dfrac{d^2y}{ds^2} & \dfrac{d^2z}{ds^2} \\ \dfrac{d^3x}{ds^3} & \dfrac{d^3y}{ds^3} & \dfrac{d^3z}{ds^3} \end{array}$$
$$\left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2 + \left(\frac{d^2z}{ds^2}\right)^2$$

In the next article a concrete example will be considered in which the flexure and torsion of a given space curve are worked out in detail.

§29

The Flexure and Torsion of a Circular Helix

In Fig. 23 is represented a right-handed circular helix. The axis of the helix is collinear with the unit vector $\mathbf{k}$ of an $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system of axes with origin at O. If $\mathbf{r}$ be the position-vector of any point P on the helix, $\mathbf{q}$ the vector projection of $\mathbf{r}$ upon the $\mathbf{i}$, $\mathbf{j}$-plane, θ the angle between $\mathbf{i}$ and $\mathbf{q}$, reckoned positive in the direction from $\mathbf{i}$ toward $\mathbf{j}$, and p the pitch of the helix, then:

(1). $$\mathbf{r} = q \cos\theta\, \mathbf{i} + q \sin\theta\, \mathbf{j} + p\,\theta\, \mathbf{k}$$

will be its vector equation; and the vector $\mathbf{q}$ can be expressed in the form:

(2) $$\mathbf{q} = q \cos\theta\, \mathbf{i} + q \sin\theta\, \mathbf{j}.$$

The scalar q represents the magnitude of the radius of the helix.

If s be the parameter of the curve representing the distance of P measured along the curve from a fiducial point P_0 on the $\mathbf{i}$-axis,

and if $d\mathbf{r}$ be the increment in $\mathbf{r}$ corresponding to the increment ds of this parameter, then:

$$d\mathbf{r} = (-q \sin \theta\, \mathbf{i} + q \cos \theta\, \mathbf{j} + p\mathbf{k})\, d\theta,$$

$$ds = \sqrt{d\mathbf{r} \cdot d\mathbf{r}} = \sqrt{q^2 + p^2} d\theta = \frac{1}{h}\, d\theta,$$

where:

$$\frac{1}{\sqrt{q^2 + p^2}} = \text{constant}. \tag{3}$$

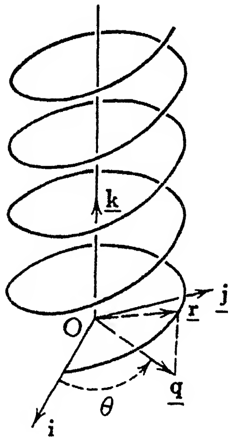


Fig. 23.

By successive differentiation of both members of equation (1) with respect to the curve parameter s, we get:

$$\frac{d\mathbf{r}}{ds} = h(-q \sin \theta\, \mathbf{i} + q \cos \theta\, \mathbf{j} + p\mathbf{k}),$$

$$\frac{d^2\mathbf{r}}{ds^2} = h^2(-q \cos \theta\, \mathbf{i} - q \sin \theta\, \mathbf{j}) = -h^2\mathbf{q}, \tag{4}$$

$$\frac{d^3\mathbf{r}}{ds^3} = h^3(q \sin \theta\, \mathbf{i} - q \cos \theta\, \mathbf{j}).$$

With the aid of these equations the flexure and the torsion of the helix can easily be calculated from equations (14) and (15), respectively, of Art. 28, with the following results:

$$C_1 = \frac{}{q^2 + p^2}, \tag{5}$$

$$C_2 = \frac{p}{q^2 + p^2}. \tag{6}$$

The flexure and torsion of the helix are positive in accordance with the conventions adopted above.

For $p = 0$ the helix degenerates into a circle in the **i**, **j**-plane for which the flexure is equal to the reciprocal of its radius, and for which the torsion vanishes.

§30
Velocity and Acceleration of a Moving Particle

Referring to Fig. 22, let the point P, whose position-vector is $\mathbf{r}$, denote the position at any time of a particle supposed moving along the curve in space there represented. For the velocity-vector and the acceleration-vector of the particle we have:

$$\text{(1)} \qquad \begin{aligned} \mathbf{v} &= \frac{d\mathbf{r}}{dt} = \frac{ds}{dt}\frac{d\mathbf{r}}{ds} = \frac{ds}{dt}\mathbf{t}, \\ \mathfrak{a} &= \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \frac{d^2s}{dt^2}\mathbf{t} + \frac{ds}{dt}\frac{d\mathbf{t}}{ds}, \end{aligned}$$

where t denotes the time parameter. But we can write:

$$\frac{d\mathbf{t}}{dt} = \frac{ds}{dt}\frac{d\mathbf{t}}{ds} = C_1\frac{ds}{dt}\mathbf{c},$$

where C_1 is the flexure of the path of the particle at P.[1] Hence:

$$\text{(2)} \qquad \mathfrak{a} = \frac{d^2s}{dt^2}\mathbf{t} + C_1\left(\frac{ds}{dt}\right)^2\mathbf{c},$$

where the first term on the right is a vector which specifies the component of the acceleration of the particle along its path due to its change of speed, and the second term is a vector which specifies the component directed toward the center of curvature of the path and proportional to the square of the speed and the flexure of the path.

It should be noticed that the velocity and acceleration-vectors both lie in the osculating plane of the path.

§31
Elements of the Theory of Surfaces

It is proposed to give in the present article a short presentation of the elements of the theory of surfaces which will enable the reader to form some idea of the usefulness of vector methods in

[1] In this connection see the foot-note in Art. 28.

connection with the differential geometry of surfaces.[1] In particular, the important subject of curvature of surfaces will receive some attention, and results will be found which will be of use subsequently.

It will prove convenient to introduce at the beginning of our brief discussion of this subject special parametric surface co-ordinates, commonly called Gaussian Co-ordinates after the German mathematician Karl Friedrich Gauss (1777–1855), whose contributions to the theory of surfaces were of fundamental importance and far reaching significance.

(a) **Gaussian surface co-ordinates.** Let x^1, x^2, x^3 denote orthogonal Cartesian co-ordinates of any point on a surface S, and suppose that to every point on the surface there corresponds a pair of values of two parameters u^1 and u^2; and vice versa, that to every pair of values of u^1 and u^2, within limits depending on the boundary of S, there corresponds a single point on the surface. Then, as in analytic geometry, the surface can be represented by the parametric equations:[2]

$$\begin{aligned} x^1 &= x^1(u^1, u^2), \\ x^2 &= x^2(u^1, u^2), \\ x^3 &= x^3(u^1, u^2). \end{aligned} \tag{1}$$

It is to be understood that the surface is regular in the sense that the functions x^1, x^2, x^3 in equations (1), together with their derivatives with respect to the parameters u^1 and u^2 to any order required in the discussion of the metrical properties of the surface, may be considered as continuous.

In order that equations (1) may actually represent the surface S, any two of them must be capable of solution for the parameters u^1 and u^2, and the values for u^1, u^2 thus obtained, when substituted in the third of these equations, must furnish an equation in the variables x^1, x^2, x^3, in fact, the equation of the surface.

If either one of the parameters, say u^1, is supposed held constant, then equations (1) together with the additional equation, $u^1 = \text{const}$, will represent a line on the surface S. Such lines are called Co-ordinate Lines. Through any point P of the surface there will pass two co-ordinate lines, as shown in Fig. 24; one of these, along which u^2 is constant and u^1 varies, is called the u^1-curve; the other, along

[1] Readers who may wish to go further into this subject are referred to the treatise of W. Blaschke-Differentialgeometrie, Band I; or to the treatise of Lagally—Vektor Rechnung, §5.

[2] The use of superscripts as identifying indices in these equations conforms to usage which will be justified later.

which u^1 is constant and u^2 varies, is called the u^2-curve. The parameters u^1 and u^2 may be considered as surface co-ordinates of the point P; parametric co-ordinates of this sort are called Gaussian Co-ordinates.

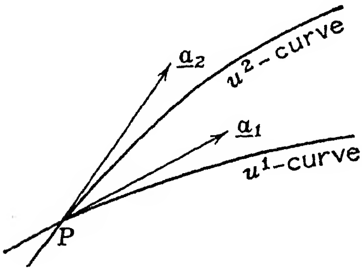


Fig. 24.

If r denote the position-vector of the point P with respect to any arbitrary origin, then, since P is a surface point, r can be considered as a function of the parameters u^1, u^2 so that:

$$\mathbf{r} = \mathbf{r}(u^1, u^2$$

By differentiation:

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u^1}\, du^1 + \frac{\partial \mathbf{r}}{\partial u^2}\, du^2, \tag{2}$$

where $d\mathbf{r}$ denotes the differential increment in r which occurs in passing from the point $P(u^1, u^2)$ to any infinitely near surface point $P'(u^1 + du^1, u^2 + du^2)$. The vectors $\partial\mathbf{r}/\partial u^1$ and $\partial\mathbf{r}/\partial u^2$ are vectors directed tangentially to the u^1 and u^2-curves, respectively, in the directions of u^1 and u^2 increasing.

These vectors will be called Unitary Vectors, and will be denoted by $\mathfrak{a}_1$ and $\mathfrak{a}_2$ respectively, so that:

$$\mathfrak{a}_1 = \frac{\partial \mathbf{r}}{\partial u^1}, \qquad \mathfrak{a}_2 = \frac{\partial \mathbf{r}}{\partial u^2}. \tag{3}$$

The unitary vectors may be considered as constituting an $\mathfrak{a}_1$, $\mathfrak{a}_2$-base-system at the point P.

Any vector associated with the point P which can be represented by a line-vector tangent to the surface at P will be called a Surface-Vector. Such a vector (A) can be expressed on the $\mathfrak{a}_1$; $\mathfrak{a}_2$-base-system by writing:

$$\mathbf{A} = A^1\mathfrak{a}_1 + A^2\mathfrak{a}_2, \tag{4}$$

where, for reasons that will appear later, A^1 and A^2 are called the Scalar Contravariant Components of **A**. As a special example we have from equation (2) for the infinitesimal surface vector $d\mathbf{r}$ the expression:

$$d\mathbf{r} = \mathfrak{a}_1 du^1 + \mathfrak{a}_2 du^2. \tag{2'}$$

(b) The first differential quadratic form. By forming the scalar product of both members of the last equation, we find:

$$\overline{ds}^2 = g_{11}du^1du^1 + 2g_{12}du^1du^2 + g_{22}du^2du^2, \tag{5}$$

where ds denotes the magnitude of $d\mathbf{r}$ and:

$$(6) \qquad g_{11} = \mathbf{a}_1 \cdot \mathbf{a}_1, \qquad g_{12} = g_{21} = \mathbf{a}_1 \cdot \mathbf{a}_2, \qquad g_{22} = \mathbf{a}_2 \cdot \mathbf{a}_2.$$

From the first and last of these equations the magnitudes of the unitary vectors $\mathbf{a}_1$, $\mathbf{a}_2$ are expressed by $\sqrt{g_{11}}$, $\sqrt{g_{22}}$, respectively.

The differential quadratic form (5) is called the First Differential Quadratic Form for the surface. The full significance and importance of this form was first recognized by Gauss, who pointed out that all the metrical properties of any surface figure can be expressed in terms of its coefficients.

If $d\mathbf{r}$ be taken in turn in the directions of u^1, u^2 increasing along the u^1, u^2-curves, and if $ds_{(1)}$, $ds_{(2)}$ denote the corresponding magnitudes of $d\mathbf{r}$ when so taken, it follows from equation (5) that:

$$(7) \qquad ds_{(1)} = \sqrt{g_{11}}\, du^1, \qquad ds_{(2)} = \sqrt{g_{22}}\, du^2.$$

The cosine of the angle between the unitary vectors at P is given, with the aid of equations (6), by the equation:

$$(8) \qquad \cos(\mathbf{a}_1, \mathbf{a}_2) = \frac{g_{12}}{\sqrt{g_{11}g_{22}}}$$

If, therefore, the coefficient g_{12} vanishes over the surface, the u^1 and u^2-curves will form an Orthogonal System of parametric lines.

If $d\mathbf{s}$ and $\delta\mathbf{s}$ denote any two infinitesimal surface vectors associated with the point P, so that:

$$(9) \qquad d\mathbf{s} = du^1\mathbf{a}_1 + du^2\mathbf{a}_2, \qquad \delta\mathbf{s} = \delta u^1\mathbf{a}_1 + \delta u^2\mathbf{a}_2;$$

then:

$$(10) \quad \cos(d\mathbf{s}, \delta\mathbf{s}) = \frac{d\mathbf{s}\cdot\delta\mathbf{s}}{ds\delta s} = \frac{g_{11}du^1\delta u^1 + g_{12}(du^1\delta u^2 + du^2\delta u^1) + g_{22}du^2\delta u^2}{ds\delta s}$$

where ds, δs denote the magnitudes of $d\mathbf{s}$, $\delta\mathbf{s}$.

For the magnitude ($d\sigma$) of the area of the infinitesimal parallelogram constructed upon line-vector elements representing $d\mathbf{s}$ and $\delta\mathbf{s}$ as sides we have:

$$\begin{aligned} d\sigma &= |d\mathbf{s} \times \delta\mathbf{s}| \\ &= |(du^1\mathbf{a}_1 + du^2\mathbf{a}_2) \times (\delta u^1\mathbf{a}_1 + \delta u^2\mathbf{a}_2)| \\ &= |\mathbf{a}_1 \times \mathbf{a}_2(du^1\delta u^2 - du^2\delta u^1)| \\ &= |a_1a_2 \sin(\mathbf{a}_1, \mathbf{a}_2)(du^1\delta u^2 - du^2\delta u^1)|, \end{aligned}$$

and hence, taking account of equations (6) and (8):

$$(11) \qquad d\sigma = \left|\sqrt{g_{11}g_{22} - g_{12}^2}\,(du^1\delta u^2 - du^2\delta u^1)\right|.$$

In the special case that ds, δs coincide in direction with the u^1, u^2-curves respectively we have $du^2 = 0$, $\delta u^1 = 0$, and hence:

$$d\sigma = \sqrt{g_{11}g_{22} - g_{12}^2}\, du^1 \delta u^2. \tag{12}$$

It is important to observe that the form (5) is not only an invariant as regards choice of co-ordinate system, but that it is also invariant to any possible distortion of the surface which does not involve stretching or tearing.

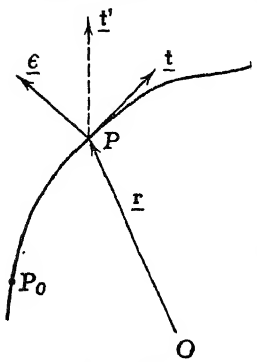


Fig. 25.

(c) **Surface curves—geodetics.** In discussing the properties of a curve which lies on a surface it is convenient to use a different reference system of unit vectors than that introduced in Art. 28 for space curves.

We therefore introduce an orthogonal right-handed system of unit vectors $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$ (see Fig. 25) associated with any point P of a curve on a surface S, the vector $\mathbf{t}$ being tangent to the curve in the direction of increase of a curve parameter s which specifies the distance of P from a fiducial point P_0 measured along the curve, the vector $\mathbf{t}'$ being perpendicular to $\mathbf{t}$ in the tangent plane to the surface at P, and the vector $\boldsymbol{\varepsilon}$ being in the direction such that $\boldsymbol{\varepsilon} = \mathbf{t} \times \mathbf{t}'$.

In passing along the curve from the point $P(s)$ to the neighboring point $P'(s + ds)$, the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$-system of unit vectors may be considered to undergo an infinitesimal translation and rotation which brings it into the configuration appropriate to the point P', and if $\mathbf{q}$ denote any vector at P which is fixed with respect to the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$-system, it will experience a change $d\mathbf{q}$ due to the rotation only which, by analogy with equation (4), Art. 28, must be such that:

$$\frac{d\mathbf{q}}{ds} = \boldsymbol{\Omega}' \times \mathbf{q}, \tag{13}$$

where, by analogy with equation (5) of the same article:

$$\boldsymbol{\Omega}' = \boldsymbol{\varepsilon} \cdot \frac{d\mathbf{t}'}{ds}\,\mathbf{t} + \mathbf{t} \cdot \frac{d\boldsymbol{\varepsilon}}{ds}\,\mathbf{t}' + \mathbf{t}' \cdot \frac{d\mathbf{t}}{ds}\,\boldsymbol{\varepsilon}. \tag{14}$$

The coefficients of the unit vectors $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$ in this equation will be denoted by T, N, G, so that:

$$T = \boldsymbol{\varepsilon}\cdot\frac{d\mathbf{t}'}{ds},$$

(15)
$$N = \mathbf{t}\cdot\frac{d\boldsymbol{\varepsilon}}{ds},$$

$$G = \mathbf{t}'\cdot\frac{d\mathbf{t}}{ds};$$

and:

(16)
$$\boldsymbol{\Omega}' = T\mathbf{t} + N\mathbf{t}' + G\boldsymbol{\varepsilon}.$$

By taking $\mathbf{q}$ in equation (13) as $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$ in turn, we get:

(17)
$$\begin{aligned}
\frac{d\mathbf{t}}{ds} &= \boldsymbol{\Omega}'\times\mathbf{t} = 0\mathbf{t} + G\mathbf{t}' - N\boldsymbol{\varepsilon},\\
\frac{d\mathbf{t}'}{ds} &= \boldsymbol{\Omega}'\times\mathbf{t}' = -G\mathbf{t} + 0\mathbf{t}' + T\boldsymbol{\varepsilon},\\
\frac{d\boldsymbol{\varepsilon}}{ds} &= \boldsymbol{\Omega}'\times\boldsymbol{\varepsilon} = N\mathbf{t} - T\mathbf{t}' + 0\boldsymbol{\varepsilon}.
\end{aligned}$$

The quantities T, N, G, respectively, are called the Geodetic Torsion, Normal Curvature, and Geodetic Curvature of the surface curve; T represents the angle of rotation of the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$- system about the $\mathbf{t}$-axis per unit distance traversed along the curve; N is numerically equal to the ordinary flexure of the projection of the curve upon the $\mathbf{t}$, $\boldsymbol{\varepsilon}$-plane, and G to the ordinary flexure of the projection of the curve upon the $\mathbf{t}$, $\mathbf{t}'$-plane. A few words relating to the signs of the quantities T, N, G as defined by equations (15) are necessary. Except in the case of a closed surface, the choice of direction for the unit surface normal $\boldsymbol{\varepsilon}$ is indeterminate, and consequently that of the unit vector $\mathbf{t}'$ is also indeterminate. The quantities N and G are therefore ambiguous as regards sign, while T is not.

A geodetic line on a surface can be defined as a curve on the surface such that the length of an arc of the curve between any two points upon it is an extremum as regards the lengths of infinitely near curves connecting the same two points. In Art. 97 the differential equations of a geodetic line on a surface will be derived. At present the discussion will be limited to showing that the necessary and sufficient condition that a curve on a surface shall be a geodetic line is that its geodetic curvature shall vanish at all points.

Consider a curve drawn on a surface S between any two points P_1 and P_2. Let s be a parameter representing the distance measured along the curve from P_1 to any point $P(s)$ between P_1 and P_2, and let $\mathbf{r}$ be the position-vector of $P(s)$.

In the special case that $d\mathbf{s}$, $\delta\mathbf{s}$ coincide in direction with the u^1, u^2-curves respectively we have $du^2 = 0$, $\delta u^1 = 0$, and hence:

$$d\sigma = \sqrt{g_{11}g_{22} - } \tag{12}$$

It is important to observe that the form (5) is not only an invariant as regards choice of co-ordinate system, but that it is also invariant to any possible distortion of the surface which does not involve stretching or tearing.

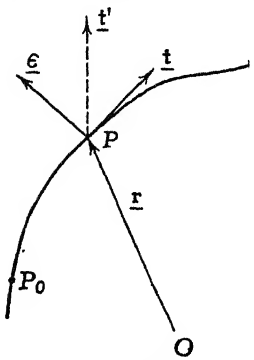


Fig. 25.

(c) **Surface curves—geodetics.** In discussing the properties of a curve which lies on a surface it is convenient to use a different reference system of unit vectors than that introduced in Art. 28 for space curves.

We therefore introduce an orthogonal right-handed system of unit vectors $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$ (see Fig. 25) associated with any point P of a curve on a surface S, the vector $\mathbf{t}$ being tangent to the curve in the direction of increase of a curve parameter s which specifies the distance of P from a fiducial point P_0 measured along the curve, the vector $\mathbf{t}'$ being perpendicular to $\mathbf{t}$ in the tangent plane to the surface at P, and the vector $\boldsymbol{\varepsilon}$ being in the direction such that $\boldsymbol{\varepsilon} = \mathbf{t} \times \mathbf{t}'$.

In passing along the curve from the point $P(s)$ to the neighboring point $P'(s + ds)$, the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$-system of unit vectors may be considered to undergo an infinitesimal translation and rotation which brings it into the configuration appropriate to the point P', and if $\mathbf{q}$ denote any vector at P which is fixed with respect to the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$-system, it will experience a change $d\mathbf{q}$ due to the rotation only which, by analogy with equation (4), Art. 28, must be such that:

$$\frac{d\mathbf{q}}{ds} = \boldsymbol{\Omega}' \times \mathbf{q}, \tag{13}$$

where, by analogy with equation (5) of the same article:

$$\boldsymbol{\Omega}' = \boldsymbol{\varepsilon} \cdot \frac{d\mathbf{t}'}{ds}\,\mathbf{t} + \mathbf{t} \cdot \frac{d\boldsymbol{\varepsilon}}{ds}\,\mathbf{t}' + \mathbf{t}' \cdot \frac{d\mathbf{t}}{ds}\,\boldsymbol{\varepsilon}. \tag{14}$$

The coefficients of the unit vectors $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$ in this equation will be denoted by T, N, G, so that:

$$
(15) \qquad \begin{aligned}
T &= \boldsymbol{\varepsilon} \cdot \frac{d\mathbf{t}'}{ds}, \\
N &= \mathbf{t} \cdot \frac{d\boldsymbol{\varepsilon}}{ds}, \\
G &= \mathbf{t}' \cdot \frac{d\mathbf{t}}{ds};
\end{aligned}
$$

and:

$$
(16) \qquad \boldsymbol{\Omega}' = T\mathbf{t} + N\mathbf{t}' + G\boldsymbol{\varepsilon}.
$$

By taking $\mathbf{q}$ in equation (13) as $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$ in turn, we get:

$$
(17) \qquad \begin{aligned}
\frac{d\mathbf{t}}{ds} &= \boldsymbol{\Omega}' \times \mathbf{t} = 0\mathbf{t} + G\mathbf{t}' - N\boldsymbol{\varepsilon}, \\
\frac{d\mathbf{t}'}{ds} &= \boldsymbol{\Omega}' \times \mathbf{t}' = -G\mathbf{t} + 0\mathbf{t}' + T\boldsymbol{\varepsilon}, \\
\frac{d\boldsymbol{\varepsilon}}{ds} &= \boldsymbol{\Omega}' \times \boldsymbol{\varepsilon} = N\mathbf{t} - T\mathbf{t}' + 0\boldsymbol{\varepsilon}.
\end{aligned}
$$

The quantities T, N, G, respectively, are called the Geodetic Torsion, Normal Curvature, and Geodetic Curvature of the surface curve; T represents the angle of rotation of the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}$- system about the $\mathbf{t}$-axis per unit distance traversed along the curve; N is numerically equal to the ordinary flexure of the projection of the curve upon the $\mathbf{t}$, $\boldsymbol{\varepsilon}$-plane, and G to the ordinary flexure of the projection of the curve upon the $\mathbf{t}$, $\mathbf{t}'$-plane. A few words relating to the signs of the quantities T, N, G as defined by equations (15) are necessary. Except in the case of a closed surface, the choice of direction for the unit surface normal $\boldsymbol{\varepsilon}$ is indeterminate, and consequently that of the unit vector $\mathbf{t}'$ is also indeterminate. The quantities N and G are therefore ambiguous as regards sign, while T is not.

A geodetic line on a surface can be defined as a curve on the surface such that the length of an arc of the curve between any two points upon it is an extremum as regards the lengths of infinitely near curves connecting the same two points. In Art. 97 the differential equations of a geodetic line on a surface will be derived. At present the discussion will be limited to showing that the necessary and sufficient condition that a curve on a surface shall be a geodetic line is that its geodetic curvature shall vanish at all points.

Consider a curve drawn on a surface S between any two points P_1 and P_2. Let s be a parameter representing the distance measured along the curve from P_1 to any point $P(s)$ between P_1 and P_2, and let $\mathbf{r}$ be the position-vector of $P(s)$.

Now suppose the points of the curve to undergo continuous infinitesimal displacements on the surface, perpendicular to the curve, the points P_1 and P_2 being held fixed, however, so that a new curve connecting P_1 and P_2 is generated upon the surface. The points $P(s)$ and $Q(s + ds)$, respectively, are supposed displaced to the points $P'(s')$ and $Q'(s' + ds')$, s' being the parameter of the new curve corresponding to the parameter s of the original curve. The vector $d\mathbf{r}$ of the original curve goes in the displacement into a vector $d\mathbf{r}'$ of the new curve.

If l denote the length of the original curve between the points P_1 and P_2 and l' that of the new curve, then:

$$l = \int_{P_1}^{P_2} ds = \int_{P_1}^{P_2} \sqrt{d\mathbf{r} \cdot d\mathbf{r}}, \qquad l' = \int_{P_1}^{P_2} ds' = \int_{P_1}^{P_2} \sqrt{d\mathbf{r}' \cdot d\mathbf{r}'};$$

and if the symbol δ denote variations due to the displacements of quantities appertaining to the original curve, then:

$$d\mathbf{r}' = d\mathbf{r} + \delta d\mathbf{r}; \qquad \delta l = l' - l.$$

Hence:

$$\begin{aligned}\delta l &= \int_{P_1}^{P_2} \sqrt{(d\mathbf{r} + \delta d\mathbf{r}) \cdot (d\mathbf{r} + \delta d\mathbf{r})} - \int_{P_1}^{P_2} \sqrt{d\mathbf{r} \cdot d\mathbf{r}} \\ &= \int_{P_1}^{P_2} (\sqrt{d\mathbf{r} \cdot d\mathbf{r} + 2d\mathbf{r} \cdot \delta d\mathbf{r}} - \sqrt{d\mathbf{r} \cdot d\mathbf{r}}),\end{aligned}$$

to quantities of the second order in the small quantity $\delta d\mathbf{r}$, and upon expanding by the binomial theorem the first radical in the last integral in terms of powers of $\delta d\mathbf{r}$, and neglecting terms of the second and higher powers in this quantity, we get:

$$\delta l \quad \int_{P_1}^{P_2} \frac{d\mathbf{r} \cdot \delta d\mathbf{r}}{\sqrt{d\mathbf{r} \cdot d\mathbf{r}}} = \int_{P_1}^{P_2} \frac{d\mathbf{r}}{ds} \cdot \delta d\mathbf{r},$$

or:

$$\delta l = \int_{P_1}^{P_2} \mathbf{t} \cdot \delta d\mathbf{r},$$

where $\mathbf{t}$ is a unit vector of the $\mathbf{t}$, $\mathbf{t}'$, $\boldsymbol{\varepsilon}'$-system at P, tangent to the curve in the direction of increasing s. This expression can be put in a form involving the geodetic curvature as follows: For the displacement of the point P, which must be collinear with the unit vector $\mathbf{t}'$, write $\delta n\mathbf{t}'$, where δn varies, of course, from point to point of the curve, and is, therefore, to be considered as a function of the parameter s. The displacement of the point $Q(s + ds)$ will then be expressed by

$$\delta n\mathbf{t}' + d(\delta n\mathbf{t}').$$

Consequently:

$$\delta d\mathbf{r} = d(\delta n\mathbf{t}') = \delta n d\mathbf{t}' + d\delta n\mathbf{t}',$$

and

$$\delta l = \int_{P_1}^{P_2} (\mathbf{t} \cdot d\mathbf{t}'\delta n + \mathbf{t} \cdot \mathbf{t}' d\delta n),$$

or, since $\mathbf{t}$ and $\mathbf{t}'$ are mutually perpendicular:

$$\delta l = -\int_{P_1}^{P_2} \mathbf{t}' \cdot \frac{d\mathbf{t}}{ds} \delta n ds,$$

or, in virtue of the third of equations (15):

$$\delta l = -\int_{P_1}^{P_2} G\delta n ds, \tag{18}$$

where G denotes the geodetic curvature of the curve at the point P.

If the original surface curve connecting the points P_1 and P_2 be a geodetic line, then, in accordance with the definition of such a line given above, δl must vanish for arbitrary values of δn, and consequently $G = 0$ at all points on a geodetic line; conversely, if for a surface curve connecting the points P_1 and P_2, $G = 0$ at all points, then $\delta l = 0$. Therefore, the vanishing of the geodetic curvature at all points of a surface curve is a necessary and sufficient condition that the curve be a geodetic line.

(d) The second differential quadratic form. We consider any surface curve passing through the point P whose position-vector with respect to any arbitrary origin is $\mathbf{r}$. As in Section (c) of the present article, let $\boldsymbol{\varepsilon}$ denote a unit vector whose direction is normal to the surface at P, and which, in fact, can be expressed as follows:

$$\boldsymbol{\varepsilon} = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{|\mathbf{a}_1 \times \mathbf{a}_2|}. \tag{19}$$

In passing along the curve from the point $P(u^1, u^2)$ to the infinitely near point $P'(u^1 + du^1, u^2 + du^2)$, suppose $\boldsymbol{\varepsilon}$ and $\mathbf{r}$ to increase by the amounts $d\boldsymbol{\varepsilon}$ and $d\mathbf{r}$. Then:

$$d\boldsymbol{\varepsilon} = \frac{\partial \boldsymbol{\varepsilon}}{\partial u^1} du^1 + \frac{\partial \boldsymbol{\varepsilon}}{\partial u^2} du^2, \qquad d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u^1} du^1 + \frac{\partial \mathbf{r}}{\partial u^2} du^2. \tag{20}$$

Upon forming the scalar product of $d\boldsymbol{\varepsilon}$ and $d\mathbf{r}$, we get:

$$d\boldsymbol{\varepsilon} \cdot d\mathbf{r} = l du^1 du^1 + 2m du^1 du^2 + n du^2 du^2, \tag{21}$$

where:

$$l = \frac{\partial \boldsymbol{\varepsilon}}{\partial u^1} \cdot \frac{\partial \mathbf{r}}{\partial u^1}, \quad m = \frac{1}{2}\left[\frac{\partial \boldsymbol{\varepsilon}}{\partial u^1} \cdot \frac{\partial \mathbf{r}}{\partial u^2} + \frac{\partial \boldsymbol{\varepsilon}}{\partial u^2} \cdot \frac{\partial \mathbf{r}}{\partial u^1}\right], \quad n = \frac{\partial \boldsymbol{\varepsilon}}{\partial u^2} \cdot \frac{\partial \mathbf{r}}{\partial u^2}. \tag{22}$$

The expression on the right of equation (21) is called the Second Differential Quadratic Form for the surface, the coefficients of the form being expressed by equations (22). These coefficients can be evaluated as follows:

Since $\boldsymbol{\varepsilon}$ is normal to the surface, and since the vectors $\partial \mathbf{r}/\partial u^1$ and $\partial \mathbf{r}/\partial u^2$ are directed tangentially to the surface, therefore:

$$\boldsymbol{\varepsilon} \cdot \frac{\partial \mathbf{r}}{\partial u^1} = 0, \qquad \boldsymbol{\varepsilon} \cdot \frac{\partial \mathbf{r}}{\partial u^2} = 0,$$

and upon partial differentiation of these expressions, the following relations are found:

$$\frac{\partial \boldsymbol{\varepsilon}}{\partial u^1} \cdot \frac{\partial \mathbf{r}}{\partial u^1} = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^1 \partial u^1}, \qquad \frac{\partial \boldsymbol{\varepsilon}}{\partial u^2} \cdot \frac{\partial \mathbf{r}}{\partial u^2} = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^2 \partial u^2},$$

$$\frac{\partial \boldsymbol{\varepsilon}}{\partial u^2} \cdot \frac{\partial \mathbf{r}}{\partial u^1} = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^2 \partial u^1}, \qquad \frac{\partial \boldsymbol{\varepsilon}}{\partial u^1} \cdot \frac{\partial \mathbf{r}}{\partial u^2} = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^1 \partial u^2},$$

and hence, from equations (22), we have:

$$l = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^1 \partial u^1},$$

$$m = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^1 \partial u^2},$$

$$n = -\boldsymbol{\varepsilon} \cdot \frac{\partial^2 \mathbf{r}}{\partial u^2 \partial u^2},$$

but by equation (19), taking account of equations (3) and (8), $\boldsymbol{\varepsilon}$ can be expressed in the form:

$$\boldsymbol{\varepsilon} = \frac{\dfrac{\partial \mathbf{r}}{\partial u^1} \times \dfrac{\partial \mathbf{r}}{\partial u^2}}{\sqrt{g_{11} g_{22} - g_{12}^2}}, \tag{23}$$

and consequently we have:

$$\frac{\left[\dfrac{\partial \mathbf{r}}{\partial u^1} \dfrac{\partial \mathbf{r}}{\partial u^2} \dfrac{\partial^2 \mathbf{r}}{\partial u^1 \partial u^1}\right]}{\sqrt{g_{11} g_{22} - g_{12}^2}},$$

$$m = -\frac{\left[\dfrac{\partial \mathbf{r}}{\partial u^1} \dfrac{\partial \mathbf{r}}{\partial u^2} \dfrac{\partial^2 \mathbf{r}}{\partial u^1 \partial u^2}\right]}{\sqrt{g_{11} g_{22} - g_{12}^2}}; \tag{24}$$

$$\frac{\left[\dfrac{\partial \mathbf{r}}{\partial u^1} \dfrac{\partial \mathbf{r}}{\partial u^2} \dfrac{\partial^2 \mathbf{r}}{\partial u^2 \partial u^2}\right]}{\sqrt{g_{11} g_{22} - g_{12}^2}}.$$

If s denote a curve parameter which is the measure of the distance along the surface curve from some fiducial point to the point P, then $d\mathbf{r}/ds$ will be a unit vector ($\mathbf{t}$) directed tangentially to the curve in the direction of s increasing. Since the vectors $\boldsymbol{\varepsilon}$ and $d\mathbf{r}/ds$ are mutually perpendicular, the equation:

$$\boldsymbol{\varepsilon} \cdot \frac{d\mathbf{r}}{ds} = 0$$

will be valid along the curve, and upon differentiation we find:

$$\boldsymbol{\varepsilon} \cdot \frac{d}{ds}\frac{d\mathbf{r}}{ds} + \frac{d\boldsymbol{\varepsilon}}{ds} \cdot \frac{d\mathbf{r}}{ds} = 0,$$

or:

$$\boldsymbol{\varepsilon} \cdot \frac{d\mathbf{t}}{ds} + \frac{d\boldsymbol{\varepsilon}}{ds} \cdot \frac{d\mathbf{r}}{ds} = 0.$$

The first term of this equation is equal to minus the quantity defined in Section (**c**) of the present article as the normal curvature N of the curve at P, and N, with the aid of equations (21) and (5), can therefore be expressed in the form:

$$N = \frac{l du^1 du^1 + 2m du^1 du^2 + n du^2 du^2}{g_{11} du^1 du^1 + 2g_{12} du^1 du^2 + g_{22} du^2 du^2}. \tag{25}$$

This equation gives the normal curvature of any surface curve through the point P whose direction ratio is $du^1 : du^2$ in terms of the coefficients of the two fundamental differential quadratic forms for the surface.

(**e**) **The Gaussian Curvature and the Mean Curvature of a surface.** We now consider only those curves through P which are normal sections of the surface. Values of N for these curves will represent flexures of normal sections of the surface at P.

Equation (25) can be written in the form:

$$(l - Ng_{11})\, du^1 du^1 + 2(m - Ng_{12})\, du^1 du^2 + (n - Ng_{22})\, du^2 du^2 = 0, \tag{26}$$

which is quadratic in the direction ratio $du^1 : du^2$. Values of N for which this equation has a double root are called the Principal Curvatures of the surface at the point P. For a double root the discriminant of the equation must vanish, so that:

$$\begin{vmatrix} l - Ng_{11} & m - Ng_{12} \\ m - Ng_{12} & n - Ng_{22} \end{vmatrix} = 0,$$

or, upon expansion:

(27) $(g_{11}g_{22} - g_{12}^2)\, N^2 - (g_{11}n - 2g_{12}m + g_{22}l)\, N + (ln - m^2) = 0.$

If N_1 and N_2 denote the two roots of this equation, then N_1 and N_2 are the principal curvatures of the surface at the point P. By the well known relations for the coefficients and the roots of a quadratic equation we have:

$$(28) \qquad N_1N_2 = \frac{ln - m^2}{g_{11}g_{22} - g_{12}^2},$$

$$(29) \qquad N_1 + N_2 = \frac{g_{11}n - 2g_{12}m + g_{22}l}{g_{11}g_{22} - g_{12}^2}$$

If K and H represent two new quantities defined by the equations:

$$(30) \qquad K = N_1N_2,$$

$$(31) \qquad H = N_1 + N_2,$$

then K is called the Gaussian Curvature, and H the Mean Curvature of the surface at the point P.

EXERCISES ON CHAPTER II

1. If $\mathbf{r} \cdot d\mathbf{r} = 0$, show that $r =$ const.
If $\mathbf{r} \times d\mathbf{r} = 0$, show that $\mathbf{r} =$ a vector const.
If $\mathbf{r} \cdot d\mathbf{r} \times d^2\mathbf{r} = 0$, show that $\mathbf{r} \times d\mathbf{r}$ has a fixed direction, and that $\mathbf{r}$ is parallel to a fixed plane.

2. Show that:

$$d\mathbf{r}_0 = -\frac{\mathbf{r}_0 \times (\mathbf{r} \times d\mathbf{r})}{r^2},$$

where $\mathbf{r}_0$ is a unit vector in the direction of $\mathbf{r}$.

3. Show vectorially that the area of the parallelogram determined by two conjugate radii of an ellipse is the same for all pairs of such radii.

4. If $\mathbf{r}$ denote the position-vector of a point P moving in a plane with respect to an origin O taken in the plane, show that the velocity-vector of the point is expressed by the equation:

$$\frac{d\mathbf{r}}{dt} = \frac{dr}{dt}\mathbf{a}_0 + r\frac{d\theta}{dt}\mathbf{b}_0,$$

where r, θ are polar co-ordinates of P, and $\mathbf{a}_0$, $\mathbf{b}_0$ are unit vectors in the directions of r, θ increasing; show also that the acceleration-vector of the point is expressed by the equation:

$$\frac{d^2\mathbf{r}}{dt^2} = \left[\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right]\mathbf{a}_0 + \left[r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt}\right]\mathbf{b}_0.$$

5. If $\mathbf{r}$ denote the position vector with respect to an arbitrary origin O of a point P moving in space, and if r, θ, ϕ denote spherical co-ordinates of the point, show that the acceleration-vector of the point can be expressed in the form:

$$\begin{aligned}\frac{d^2\mathbf{r}}{dt^2} = &\quad [r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\dot{\phi}^2 \sin\theta\cos\theta]\,\mathbf{a}_0 \\ &+ [2\dot{r}\dot{\phi}\sin\theta + 2r\dot{\theta}\dot{\phi}\cos\theta + r\ddot{\phi}\sin\theta]\,\mathbf{b}_0 \\ &+ [\ddot{r} - r\dot{\theta}^2 - r\dot{\phi}^2\sin^2\theta]\,\mathbf{c}_0,\end{aligned}$$

where $\mathbf{a}_0$, $\mathbf{b}_0$, $\mathbf{c}_0$ are unit vectors in the directions of r, θ, ϕ increasing.

6. If $\mathbf{r}$ and $\mathbf{r} + \delta\mathbf{r}$ are position vectors of two neighboring points $P(s)$ and $P'(s + \delta s)$ on a space curve, δs being the length of the arc PP', show that:

$$\mathbf{t}\cdot\delta\mathbf{r} \text{ is of the order of magnitude } \delta s,$$

$$\mathbf{c}\cdot\delta\mathbf{r} \text{ is of the order of magnitude } \overline{\delta s}^{\,2},$$

$$\mathbf{n}\cdot\delta\mathbf{r} \text{ is of the order of magnitude } \overline{\delta s}^{\,3},$$

where $\mathbf{t}$, $\mathbf{c}$, $\mathbf{n}$ are unit vectors whose significance is explained in Art. 28.

7. Show that the solutions of the differential equations:

$$\frac{d^2\mathbf{r}}{dt^2} = \pm k^2\mathbf{r}$$

are:

$$\mathbf{r} = \mathbf{A}e^{kt} + \mathbf{B}e^{-kt},$$
$$\mathbf{r} = \mathbf{A}\cos kt + \mathbf{B}\sin kt,$$

where $\mathbf{A}$ and $\mathbf{B}$ are vector constants of integration. Find the flexures of the orbits represented by the last two equations.

8. Show that any space curve whose flexure and torsion are in a constant ratio must be a cylindrical helix.

9. The circle of curvature at a point P of a space curve is a circle which passes through P and two neighboring points on the curve when they coalesce with P. Show that the locus of the center of the circle of curvature of a circular helix is itself a circular helix.

10. Find the flexure and torsion of a helix on a cone which makes a constant angle with the elements of the cone.

11. By means of formula (30), Art. 31, show that the Gaussian curvature of a sphere is equal to the reciprocal of the square of its radius.

12. The position-vector $\mathbf{r}$ of a point P is expressed in terms of its $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-components as follows:

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Suppose that x, y, z are functions of three parameters u, v, w whose Jacobean is:

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

Give the geometrical significance of this Jacobean.

CHAPTER III

SCALAR AND VECTOR FIELDS

§32

Scalar and Vector Point Functions

The position of a point in space can be specified by three numbers expressing the values of a corresponding set of quantities (q_1, q_2, q_3) called co-ordinates; ordinary Cartesian and spherical co-ordinates are familiar special examples.

Associated with each point of space let us suppose a scalar function u and a vector function $\mathbf{v}$, the values of which at any point depend only upon the values of the co-ordinates of the point, so that:

$$u = u(q_1, q_2, q_3), \qquad \mathbf{v} = \mathbf{v}(q_1, q_2, q_3).$$

Such functions are called Scalar and Vector Point Functions respectively. In the preceding pages we have already had occasionally to deal with such functions.

Examples of scalar point functions met with in Physics are presented by: the density of a distribution of gravitational matter, the temperature in a body or system of bodies, the electrostatic potential of a distribution of electric charges; and examples of vector point functions by: the gravitational force associated with a distribution of gravitational matter, the velocity of a moving fluid and the electric field intensity due to a distribution of electric charges.

Unless otherwise stated, scalar and vector point functions together with their first space derivatives will be considered to be continuous and singly valued functions.

The ensemble of points in a given region together with the corresponding values of a scalar point function, or of a vector point function, constitute what is called the Field of the Scalar Point Function, or the Field of the Vector Point Function, in the region.

Any point in the field of a scalar or a vector point function is called a Field Point.

§33

Relating to the Field of a Scalar Point Function

In the consideration of the field of a scalar point function it is advantageous to introduce the concept of a level surface of the function:

A Level Surface of a scalar point function is a surface for all points of which the function has the same value.

Using rectangular Cartesian co-ordinates x, y, z, we can take for the equation of a level surface of a scalar point function u:

$$u(x, y, z) = c, \tag{1}$$

where c is a constant; and the position-vector $\mathbf{r}$ of the field-point $P(x, y, z)$ with which u is associated can be expressed in the form:

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}. \tag{2}$$

Referring to Fig. 26, let P be a point on the level surface $u =$ const, and P' a neighboring point on an infinitely near level surface

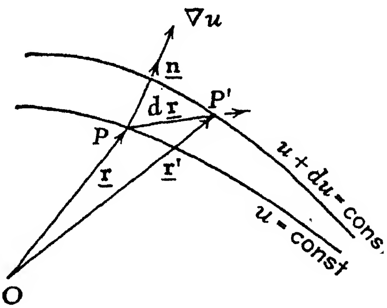


Fig. 26.

for which the value of u is greater by the amount du. Let the co-ordinates and position-vector of P' be denoted by $x + dx$, $y + dy$, $z + dz$, and $\mathbf{r}'$. Then:

$$\mathbf{r}' = (x + dx)\,\mathbf{i} + (y + dy)\,\mathbf{j} + (z + dz)\,\mathbf{k},$$

$$d\mathbf{r} = \mathbf{r}' - \mathbf{r} = \mathbf{i}dx + \mathbf{j}dy + \mathbf{k}dz,$$

$$du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy + \frac{\partial u}{\partial z}\,dz.$$

The form of the expression for du indicates that it can be written as a scalar product as follows:

$$du = (\mathbf{i}dx + \mathbf{j}dy + \mathbf{k}dz)\cdot\left(\mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z}\right). \tag{3}$$

The first vector factor on the right represents the infinitesimal vector $d\mathbf{r}$. The second represents a vector of fundamental importance in the theory of scalar fields, called the Gradient of the scalar point function u at the point P; it will, following Hamilton, be denoted by ∇u, so that:

(4) $$\nabla u = \mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z}.$$

The symbol ∇, read "del," defined by writing:

(5) $$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z},$$

is an operator which, acting upon the scalar point function u, produces the gradient of this function.

Equation (3) can now be written:

(6) $$du = d\mathbf{r} \cdot \nabla u = \nabla u \cdot d\mathbf{r}.$$

In Art. 36 a definition of the gradient of a scalar point function will be given which shows that it is an invariant as regards the co-ordinate system with respect to which it is expressed. Meanwhile, it will be convenient to call attention to a geometrical interpretation which shows the real significance of this important vector.

In Fig. 26 $\mathbf{n}$ represents a unit normal at P to the level surface of u through P in the direction of u increasing. Hence, if ds denote the magnitude of $d\mathbf{r}$, and dn the value of ds when P' is so taken that $d\mathbf{r}$ coincides in direction with $\mathbf{n}$, then:

$$dn = \mathbf{n} \cdot d\mathbf{r},$$

and:

$$du = \frac{du}{dn}dn = \frac{du}{dn}\mathbf{n} \cdot d\mathbf{r};$$

but, as seen above:

$$du = \nabla u \cdot d\mathbf{r};$$

hence:

$$\nabla u \cdot d\mathbf{r} = \frac{du}{dn}\mathbf{n} \cdot d\mathbf{r}.$$

Since this equation is valid for all possible values of $d\mathbf{r}$, it follows that:

(7) $$\nabla u = \frac{du}{dn}\mathbf{n}.$$

The gradient of u is therefore a vector in the direction of $\mathbf{n}$, which is that of the greatest space rate of increase of u, and, since $\mathbf{n}$ is a unit vector, its magnitude, viz. du/dn, is equal to the greatest space rate of increase of u.

The space rate of increase of u in the general direction indicated by a unit vector $\mathbf{s}_0$ can be expressed as follows:

$$\frac{du}{ds} = \frac{d\mathbf{r}}{ds} \cdot \nabla u = \mathbf{s}_0 \cdot \nabla u. \tag{8}$$

Owing to the fact that the space rate of increase of a scalar point function is dependent upon direction, it is often denoted by a partial derivative; thus:

$$\frac{\partial u}{\partial s} = \mathbf{s}_0 \cdot \nabla u, \tag{9}$$

but it is not, of course, hereby implied that u is an integral function of s.

As will be seen later, it is oftentimes advantageous to "map" the field of a scalar point function u by drawing successive level surfaces of u, the difference in the values of u for two successive surfaces differing by a constant, small, finite amount, and then drawing a complementary system of curves in the directions of ∇u; these curves are called Field Lines, and intersect the level surfaces orthogonally.

The gradient of u at any point will have the direction of a corresponding field line at that point, and if $d\mathbf{r}$ represent a vector element of this line, the vector equation of the field line will be:

$$d\mathbf{r} \times \nabla u = 0, \tag{10}$$

the Cartesian equivalents of which are the equations:

$$dy\frac{\partial u}{\partial z} - dz\frac{\partial u}{\partial y} = 0,$$
$$dz\frac{\partial u}{\partial x} - dx\frac{\partial u}{\partial z} = 0,$$
$$dx\frac{\partial u}{\partial y} - dy\frac{\partial u}{\partial x} = 0,$$

or:

$$dx : dy : dz = \frac{\partial u}{\partial x} : \frac{\partial u}{\partial y} : \frac{\partial u}{\partial z}. \tag{11}$$

§34

Relating to the Field of a Vector Point Function

Consider a vector point function $\mathbf{v}(x, y, z)$. At any field point $P(x, y, z)$ the corresponding value of the function can be represented by a line-vector with a length equal to the magnitude of $\mathbf{v}$ at the point, and drawn in the direction of $\mathbf{v}$. The position-vector $\mathbf{r}$ of P with respect to the arbitrary origin of co-ordinates and the function $\mathbf{v}$ can be expressed in the forms:

(1) $$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \qquad \mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}.$$

In passing from $P(x, y, z)$ to a neighboring point $P'(x + dx, y + dy, z + dz)$ with position-vector $\mathbf{r}'$, the vector $\mathbf{v}$ undergoes a change corresponding to the change in $\mathbf{r}$ expressed by:

(2) $$d\mathbf{r} = \mathbf{i}dx + \mathbf{j}dy + \mathbf{k}dz.$$

For the corresponding increment in $\mathbf{v}$ we can write:

$$\begin{aligned} d\mathbf{v} &= \frac{\partial \mathbf{v}}{\partial x}dx + \frac{\partial \mathbf{v}}{\partial y}dy + \frac{\partial \mathbf{v}}{\partial z}dz \\ &= \left(\frac{\partial v_1}{\partial x}dx + \frac{\partial v_1}{\partial y}dy + \frac{\partial v_1}{\partial z}dz\right)\mathbf{i} \\ &\quad + \left(\frac{\partial v_2}{\partial x}dx + \frac{\partial v_2}{\partial y}dy + \frac{\partial v_2}{\partial z}dz\right)\mathbf{j} \\ &\quad + \left(\frac{\partial v_3}{\partial x}dx + \frac{\partial v_3}{\partial y}dy + \frac{\partial v_3}{\partial z}dz\right)\mathbf{k}, \end{aligned}$$

or:

(3) $$d\mathbf{v} = \nabla v_1 \cdot d\mathbf{r}\mathbf{i} + \nabla v_2 \cdot d\mathbf{r}\mathbf{j} + \nabla v_3 \cdot d\mathbf{r}\mathbf{k}.$$

If ds denote the magnitude of $d\mathbf{r}$ and $\mathbf{s}_0$ a unit vector in the direction of $d\mathbf{r}$, we have for the space rate of change of $\mathbf{v}$ (directional derivative) in the direction of $\mathbf{s}_0$:

$$\frac{d\mathbf{v}}{ds} = \frac{\partial \mathbf{v}}{\partial s} = \frac{\partial \mathbf{v}}{\partial x}\frac{dx}{ds} + \frac{\partial \mathbf{v}}{\partial y}\frac{dy}{ds} + \frac{\partial \mathbf{v}}{\partial z}\frac{dz}{ds},$$

or:

(4) $$\frac{d\mathbf{v}}{ds} = \frac{\partial \mathbf{v}}{\partial s} = \mathbf{s}_0 \cdot \nabla v_1\mathbf{i} + \mathbf{s}_0 \cdot \nabla v_2\mathbf{j} + \mathbf{s}_0 \cdot \nabla v_3\mathbf{k}.$$

The vector equation of a field line of $\mathbf{v}$ is:

(5) $$d\mathbf{r} \times \mathbf{v} = 0,$$

the Cartesian equivalents of which are the equations:

(6) $$dx : dy : dz = v_1 : v_2 : v_3.$$

The character of the field of a vector point function may vary greatly in passing from point to point. Peculiarities may be present at some points which are non-existent at others. But, if the function vanishes at infinity, it will be completely determined, as will be seen later, by the specification of the values of two quantities at all points; one, a scalar, called the divergence of the vector point function, and the other, a vector, called the curl of the vector point function. These quantities represent fundamental concepts relating to the nature of a vector field, and will be defined in Art. 36.

§35

Diagrams of Fields of Scalar and Vector Point Functions

The diagrams shown in figures 27 and 28, from Maxwell's treatise on Electricity and Magnetism, are introduced to illustrate

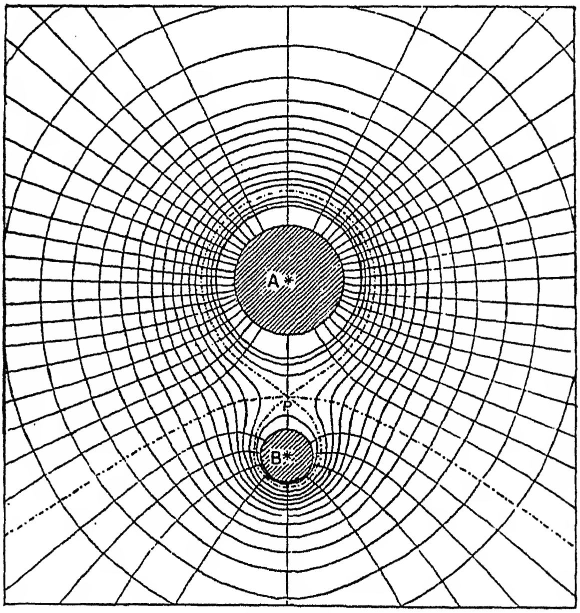


Fig. 27.

the method of "mapping" fields of scalar and vector point functions.

The diagram in Fig. 27 represents equally well the gravitational field due to two small material spheres, A and B, of the same density

whose masses are in the ratio of 4 to 1, and the electric field due to two small conducting spheres carrying electric charges of the same sign whose magnitudes are in the ratio of 4 to 1.

The field itself is symmetrical about an axial line through the center of the two spheres, and the diagram represents a section of the field in a plane through this axis.

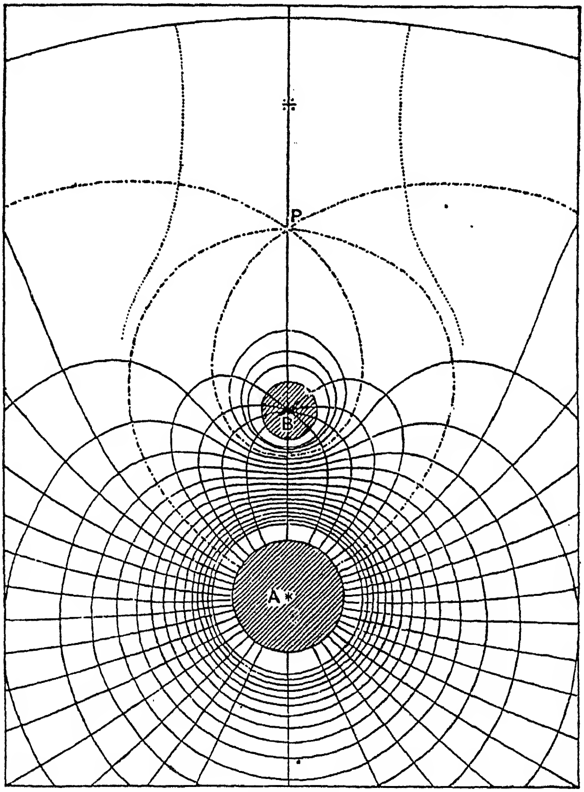


Fig. 28.

The closed lines represent level surfaces of gravitational or electric potential, and the lines perpendicular to these represent lines of gravitational or electric force which at each point is co-directional with the gradient of the corresponding potential.

The level surface, indicated by a dotted line consisting of two lobes meeting at the point P, separates the level surfaces which surround both spheres from those which surround the spheres A

and B individually. At P the resultant gravitational or electric force is zero.

The dotted line, resembling an hyperbola, represents a surface dividing the lines of gravitational or electric force into two distinct groups.

As with all diagrams of the sort here represented, the difference of potential between any two successive level surfaces is the same, and the lines of force are so drawn that their density is proportional to the force.

The diagram in Fig. 28 represents the field due to electric charges on two small spheres, A and B, the charges being of opposite signs and with magnitudes in the ratio of 4 to 1.

Here, again, the closed lines represent level or equipotential surfaces in any plane through an axial line passing through the centers of the two spheres, and the lines perpendicular to these represent lines of electric force.

The level surface, indicated by a dotted line consisting of two lobes, one surrounding the sphere B and the other surrounding both spheres, and meeting in the point P, separates the equipotential surfaces which surround both spheres and those which surround the spheres A and B individually. At the point P the resultant electric force is zero.

The lines of force are separated into two distinct groups by a surface indicated by a dotted line beginning and ending at the sphere A, and passing through the neutral point P.

For the method of construction of such diagrams the reader is referred to Maxwell's treatise, cited above, Art. 123.

§36

The Gradient of a Scalar Point Function and the Divergence and the Curl of a Vector Point Function

These quantities will be defined below in ways which differ somewhat from those usually followed by writers on vector analysis, but which have the advantage of making evident their invariance as regards co-ordinate systems.

Referring to Fig. 29, let P be a point in the field of a scalar point function u, or of a vector point function $\mathbf{v}$, these functions together with their first space derivatives being assumed finite, continuous and single-valued. Imagine a small element of volume of magnitude δ containing the point P and bounded by a small closed sur-

face ω.[1] Let the magnitude of a typical differential element of area of ω be denoted by $d\sigma$, and let $\boldsymbol{\varepsilon}$ denote a unit vector in the direction of an outward drawn normal to ω.

The gradient of the scalar point function u at the point P, denoted by grad u, is defined by the equation:[2]

(1) $$\text{grad } u = \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u \boldsymbol{\varepsilon} d\sigma.$$

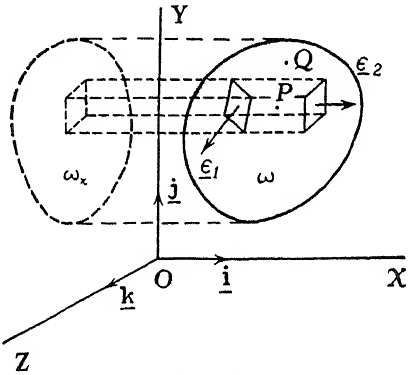


Fig. 29.

The divergence of the vector point function **v** *at the point P, denoted by div* **v**, *and the curl of the vector point function* **v**, *denoted by curl* **v**, *are respectively defined by the equations:*

(2) $$\text{div } \mathbf{v} \quad \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega \boldsymbol{\varepsilon} \cdot \mathbf{v} d\sigma;$$

(3) $$\text{curl } \mathbf{v} = \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega \boldsymbol{\varepsilon} \times \mathbf{v} d\sigma.$$

The significance of these definitions and the proof that they are really unambiguous have yet to be given.

We first fix our attention upon equation (1). Let the co-ordinates of any point Q on the surface ω be denoted by $x + f$, $y + g$, $z + h$, the co-ordinates of the point P being x, y, z. The small quantities f, g, h will thus be the relative co-ordinates of Q with respect to

[1] This surface may have a finite number of sharp edges as, for example, in the case of the surface of a cube. It will be assumed for simplicity that the surface is non-reëntrant, but this restriction is not essential.

[2] The definition of the gradient of a scalar point function here given is that mentioned in Art. 33 in connection with a previous definition.

P. If u_Q denote the value at Q of the scalar point function whose value at P is denoted by u, then, by Taylor's theorem:

$$u_Q = u + \frac{\partial u}{\partial x} f + \frac{\partial u}{\partial y} g + \frac{\partial u}{\partial z} h + \phi, \tag{4}$$

where ϕ represents all terms of higher order than the first in f, g, h, and where the partial derivatives are supposed evaluated at P. By equation (1) we have:

$$\text{grad } u = \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u\boldsymbol{\varepsilon} d\sigma = \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u(\boldsymbol{\varepsilon} \cdot \mathbf{ii} + \boldsymbol{\varepsilon} \cdot \mathbf{jj} + \boldsymbol{\varepsilon} \cdot \mathbf{kk})\, d\sigma.$$

and hence:

$$\text{grad } u = \mathbf{i} \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u\boldsymbol{\varepsilon} \cdot \mathbf{i} d\sigma + \mathbf{j} \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u\boldsymbol{\varepsilon} \cdot \mathbf{j} d\sigma + \mathbf{k} \underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u\boldsymbol{\varepsilon} \cdot \mathbf{k} d\sigma.$$

Now let $d\sigma_1$ and $d\sigma_2$ denote the magnitudes of the elements of area cut from the surface ω by an elementary prism parallel to the X-axis, and let $d\sigma_x$ denote the magnitude of the projections of these elements upon the Y-Z-plane, and ω_x that of the surface ω itself; then, if $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ denote the values of $\boldsymbol{\varepsilon}$ associated with $d\sigma_1$ and $d\sigma_2$, respectively:

$$d\sigma_x = -\boldsymbol{\varepsilon}_1 \cdot \mathbf{i} d\sigma_1 = \boldsymbol{\varepsilon}_2 \cdot \mathbf{i} d\sigma_2.$$

Making use of these relations, the integral in the first term on the right of the preceding equation can be transformed into an equivalent one for which the domain of integration is ω_x instead of ω. For, denoting quantities associated with $d\sigma_1$ and $d\sigma_2$ by subscripts 1 and 2, respectively, we can write:

$$\int_\omega u\boldsymbol{\varepsilon} \cdot \mathbf{i} d\sigma = \int_{\omega_x} (u_2 - u_1)\, d\sigma_x;$$

or, with the aid of equation (4):

$$\int_\omega u\boldsymbol{\varepsilon} \cdot \mathbf{i} d\sigma = \int_x \left\{ \frac{\partial u}{\partial x}(f_2 - f_1) + \phi_2 - \phi_1 \right\} d\sigma_x$$

$$= \frac{\partial u}{\partial x} \int_{\omega_x} \left\{ f_2 - f_1 + \frac{(\phi_2 - \phi_1)}{\partial u/\partial x} \right\} d\sigma_x.$$

Hence:

$$\underset{\delta \to 0}{Lt} \frac{1}{\delta} \int_\omega u\boldsymbol{\varepsilon} \cdot \mathbf{i} d\sigma = \frac{\partial u}{\partial x} \underset{\delta \to 0}{Lt} \left[\frac{1}{\delta} \int_x \left\{ f_2 - f_1 + \frac{(\phi_2 - \phi_1)}{\partial u/\partial x} \right\} d\sigma_x \right].$$

With the aid of Fig. 29, remembering the significance of ϕ, the limit of the expression in square brackets is seen to be unity; it follows that:

$$
\begin{aligned}
&\underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega u\boldsymbol{\varepsilon}\cdot\mathbf{i}d\sigma = \frac{\partial u}{\partial x},\\
(5)\qquad &\underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega u\boldsymbol{\varepsilon}\cdot\mathbf{j}d\sigma = \frac{\partial u}{\partial y},\\
&\underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega u\boldsymbol{\varepsilon}\cdot\mathbf{k}d\sigma = \frac{\partial u}{\partial z},
\end{aligned}
$$

the last two of these equations being obtained from the first by analogy. Hence:

$$
(6)\qquad \text{grad } u = \mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z}.
$$

This expression for the gradient of u is identical with the vector defined as the gradient of a scalar point function u in Art. 33.

Returning now to the defining equation (2) for the divergence of a vector $\mathbf{v}$, we have:

$$
\text{div } \mathbf{v} = \underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega \boldsymbol{\varepsilon}\cdot\mathbf{v}d\sigma = \underset{\delta\to 0}{Lt}\left[\frac{1}{\delta}\int_\omega v_1\boldsymbol{\varepsilon}\cdot\mathbf{i}d\sigma + \frac{1}{\delta}\int_\omega v_2\boldsymbol{\varepsilon}\cdot\mathbf{j}d\sigma + \frac{1}{\delta}\int_\omega v_3\boldsymbol{\varepsilon}\cdot\mathbf{k}d\sigma\right],
$$

where v_1, v_2, v_3 are the measure-numbers of the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-components of $\mathbf{v}$; and from equations (5), by taking u in turn as v_1, v_2, v_3, we also have:

$$
\underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega v_1\boldsymbol{\varepsilon}\cdot\mathbf{i}d\sigma = \frac{\partial v_1}{\partial x},\qquad \underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega v_2\boldsymbol{\varepsilon}\cdot\mathbf{j}d\sigma = \frac{\partial v_2}{\partial y},\qquad \underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega v_3\boldsymbol{\varepsilon}\cdot\mathbf{k}d\sigma = \frac{\partial v_3}{\partial z}.
$$

Hence:

$$
(7)\qquad \text{div } \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}.
$$

Starting now with the defining equation (3) for curl $\mathbf{v}$, we have:

$$
\begin{aligned}
\text{curl } \mathbf{v} &= \underset{\delta\to 0}{Lt}\frac{1}{\delta}\int_\omega \boldsymbol{\varepsilon}\times\mathbf{v}d\sigma\\
&= \underset{\delta\to 0}{Lt}\left[\frac{1}{\delta}\int_\omega(\boldsymbol{\varepsilon}\times\mathbf{v}\cdot\mathbf{i})\,\mathbf{i}d\sigma + \frac{1}{\delta}\int_\omega(\boldsymbol{\varepsilon}\times\mathbf{v}\cdot\mathbf{j})\,\mathbf{j}d\sigma + \frac{1}{\delta}\int_\omega(\boldsymbol{\varepsilon}\times\mathbf{v}\cdot\mathbf{k})\,\mathbf{k}d\sigma\right]\\
&= \underset{\delta\to 0}{Lt}\left[\frac{\mathbf{i}}{\delta}\int_\omega(\boldsymbol{\varepsilon}\cdot\mathbf{v}\times\mathbf{i})\,d\sigma + \frac{\mathbf{j}}{\delta}\int_\omega(\boldsymbol{\varepsilon}\cdot\mathbf{v}\times\mathbf{j})\,d\sigma + \frac{\mathbf{k}}{\delta}\int_\omega(\boldsymbol{\varepsilon}\cdot\mathbf{v}\times\mathbf{k})\,d\sigma\right]\\
&= \mathbf{i}\text{ div }\mathbf{v}\times\mathbf{i} + \mathbf{j}\text{ div }\mathbf{v}\times\mathbf{j} + \mathbf{k}\text{ div }\mathbf{v}\times\mathbf{k},
\end{aligned}
$$

with the aid of formula (2); and by formula (7) we have:

$$
\begin{aligned}
\text{div }\mathbf{v}\times\mathbf{i} &= \text{div }(v_3\mathbf{j} - v_2\mathbf{k}) = \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right),\\
\text{div }\mathbf{v}\times\mathbf{j} &= \text{div }(v_1\mathbf{k} - v_3\mathbf{i}) = \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right),\\
\text{div }\mathbf{v}\times\mathbf{k} &= \text{div }(v_2\mathbf{i} - v_1\mathbf{j}) = \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right).
\end{aligned}
$$

Hence:

$$\text{(8)} \qquad \operatorname{curl} \mathbf{v} = \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}.$$

Since in the above demonstrations no special form has been assumed for the surface ω, and since no special assumption has been made as to the manner of shrinkage of the volumes toward its zero limit, it follows that the definitions given for the gradient of a scalar point function and for the divergence and curl of a vector point function are quite unambiguous. Furthermore, since the definitions themselves involve no reference to a co-ordinate system, grad u, div $\mathbf{v}$, and curl $\mathbf{v}$ must be invariants as regards co-ordinate systems.

By taking δ in the defining equations (1), (2), and (3) equal to $d\tau$, the magnitude of a differential element of volume, the equations can be written in the following approximate forms:

$$\text{(9)} \qquad \operatorname{grad} u = \frac{1}{d\tau}\int_\omega u\boldsymbol{\varepsilon} d\sigma,$$

$$\text{(10)} \qquad \operatorname{div} \mathbf{v} = \frac{1}{d\tau}\int_\omega \boldsymbol{\varepsilon} \cdot \mathbf{v} d\sigma,$$

$$\text{(11)} \qquad \operatorname{curl} \mathbf{v} = \frac{1}{d\tau}\int_\omega \boldsymbol{\varepsilon} \times \mathbf{v} d\sigma,$$

the approximation implied being as close as we please, since $d\tau$ may be taken as small as we please. We shall return presently to further discussion of the significance of these quantities.

Meanwhile, with the aid of equations (9) and (11), we shall derive two important relations which will be found of use later.

At any point in the field of the scalar point function u, or the vector point function $\mathbf{v}$, let us consider a surface element of area of magnitude $d\sigma'$ bounded by a contour c, and let $d\mathbf{s}$ denote a differential vector directed along c in the direction around c arbitrarily chosen as positive, and $\mathbf{n}$ a unit vector normal to the element of area in the direction related to the positive direction around c as the thrust and twist of a right-handed screw. It will be shown that:

$$\text{(12)} \qquad \mathbf{n} \times \operatorname{grad} u = \frac{1}{d\sigma'}\int_c u d\mathbf{s};$$

$$\text{(13)} \qquad \mathbf{n} \cdot \operatorname{curl} \mathbf{v} = \frac{1}{d\sigma'}\int_c \mathbf{v} \cdot d\mathbf{s}.$$

Referring to equations (9) and (11), the form of the differential element of volume of magnitude $d\tau$ can be chosen arbitrarily, and it suits our purpose to take it as a small right cylinder whose top surface is the surface element in question, and whose height h is infinitesimal. See Fig. 30. Consider now the two surface integrals over the surface of the small cylinder of the quantities $u\mathbf{n} \times \boldsymbol{\varepsilon}$ and $\mathbf{n} \cdot \boldsymbol{\varepsilon} \times \mathbf{v}$. It is evident that the top and bottom surfaces will contribute nothing to these integrals, since for them $\mathbf{n} \times \boldsymbol{\varepsilon} = 0$ and $\mathbf{n} \cdot \boldsymbol{\varepsilon} \times \mathbf{v} = 0$. We find then for the integrals in question:

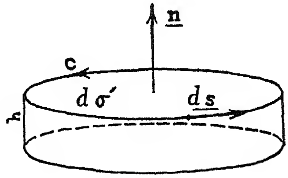


Fig. 30.

$$\int_\omega u\mathbf{n} \times \boldsymbol{\varepsilon} d\sigma = \int_c uh\mathbf{n} \times \boldsymbol{\varepsilon}|ds| = h\int_c u d\mathbf{s} = \frac{d\tau}{d\sigma'}\int_c u d\mathbf{s},$$

$$\int_\omega \mathbf{n} \cdot \boldsymbol{\varepsilon} \times \mathbf{v} d\sigma = \int_c h\mathbf{v} \cdot \mathbf{n} \times \boldsymbol{\varepsilon}|ds| = h\int_c \mathbf{v} \cdot d\mathbf{s} = \frac{d\tau}{d\sigma'}\int_c \mathbf{v} \cdot d\mathbf{s},$$

and hence, from equations (9) and (11):

$$\mathbf{n} \times \operatorname{grad} u = \frac{1}{d\tau}\int_\omega u\mathbf{n} \times \boldsymbol{\varepsilon} d\sigma = \frac{1}{d\sigma'}\int_c u d\mathbf{s},$$

$$\mathbf{n} \cdot \operatorname{curl} \mathbf{v} = \frac{1}{d\tau}\int_\omega \mathbf{n} \cdot \boldsymbol{\varepsilon} \times \mathbf{v} d\sigma = \frac{1}{d\sigma'}\int_c \mathbf{v} \cdot d\mathbf{s}.$$

The validity of equations (12) and (13) is thus established.

The significance of the gradient ∇u of a scalar point function u has already been discussed in Art. 33, where it was shown that ∇u is a vector whose magnitude and direction are those of the greatest space rate of increase of the function.

Some light is thrown upon the significance of the divergence and of the curl of a vector point function $\mathbf{v}$ upon closer consideration of equations (10) and (13).

From equation (10) it appears that, if div $\mathbf{v}$ is to have a value different from zero at any point P, the vector $\mathbf{v}$ itself must behave in such a way in the neighborhood of the point that the surface integral of its normal component over the surface bounding a differential element of volume of magnitude $d\tau$, including the point P, shall not vanish in virtue of cancellation of plus and minus contributions from different parts of the surface. For example: let us consider the electric field due to a positive charge of electricity with volume density ρ supposed distributed uniformly throughout the volume (Δ) of a small sphere with infinitesimal

radius ϵ and with center at P. If $\mathbf{E}$ denote electric field intensity, then the field lines of $\mathbf{E}$ will be directed radially outward at all points on the surface of the sphere, and $E(= \mathbf{n} \cdot \mathbf{E})$ will be uniform over the surface. Consequently, the surface integral of $\mathbf{n} \cdot \mathbf{E}$ cannot vanish in virtue of cancellation of plus and minus contributions to the integral; in fact, the contributions from all parts of the surface will be positive, and in total amount equal to $4\pi\epsilon^2 E$; therefore, by equation (10), we shall have for the divergence of $\mathbf{E}$ at P:

$$\operatorname{div} \mathbf{E} = \frac{4\pi\epsilon^2 E}{\Delta}.$$

In accordance with a fundamental electrical law, if E be supposed expressed in "rational units," then $E = \rho\Delta/4\pi\epsilon^2$, and hence:[1]

$$\operatorname{div} \mathbf{E} = \rho. \tag{14}$$

From equation (13) it appears that the scalar component of curl $\mathbf{v}$ in a direction determined by the unit vector $\mathbf{n}$ is equal to the integral of $\mathbf{v} \cdot d\mathbf{r}$ around the contour of the element of area of magnitude $d\sigma'$ (for which $\mathbf{n}$ is the unit normal) divided by $d\sigma'$. If curl $\mathbf{v}$ has a value different from zero, and if $\mathbf{n}$ be taken in the direction of curl $\mathbf{v}$, then obviously the integral of $\mathbf{v} \cdot d\mathbf{s}$ around the contour cannot vanish in virtue of cancellation of plus and minus contributions from different parts of the contour. For example: let us consider the magnetic field due to an electric current in an infinitely long straight wire, with infinitesimal section. As is well known, the lines of magnetic force ($\mathbf{H}$) in this case are circles in planes normal to the wire, and concentric with it. Now, by equation (13), if ω represent any section of the wire at any point:

$$\mathbf{n} \cdot \operatorname{curl} \mathbf{H} = \frac{1}{\omega} \int_c \mathbf{H} \cdot d\mathbf{r}.$$

The integral on the right represents the work which the magnetic field would do upon a unit positive magnetic pole passing once around the contour c bounding ω, and by a fundamental law of electromagnetism this is proportional to the current $(\mathbf{n} \cdot \mathbf{q}\omega)$ embraced by the contour, $\mathbf{q}$ representing the current per unit area; if rational units be used, the constant of proportionality is equal

[1] On the "rational" system of electrical units a unit of charge is defined as one which would repel a like charge at a distance of one centimeter with a force of $1/4\pi$ dynes; a unit magnetic pole is defined in a similar way.

to the reciprocal of the velocity of light (c) in vacuo. Hence, in these units:

$$\mathbf{n} \cdot \operatorname{curl} \mathbf{H} = \frac{1}{c} \mathbf{n} \cdot \mathbf{q}.$$

This result is valid for *any* section ω. Hence:

$$\operatorname{curl} \mathbf{H} = \frac{1}{c} \mathbf{q}. \tag{15}$$

Subsequently, from time to time, we shall meet with other examples which will serve to throw further light upon the significance of the divergence and of the curl of a vector point function.

§37
The Operator ∇ and Some Related Operators

The operator ∇ is expressed on any $\mathbf{i}, \mathbf{j}, \mathbf{k}$-base-system as follows:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}.$$

Operating upon a scalar point function u it produces, as we have seen, the gradient of the function:

$$\nabla u = \mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z}. \tag{1}$$

If $\mathbf{s}$ denote a vector point function, then:

$$\mathbf{s} \cdot \nabla u = s_1\frac{\partial u}{\partial x} + s_2\frac{\partial u}{\partial y} + s_3\frac{\partial u}{\partial z}. \tag{2}$$

where s_1, s_2, s_3 are the measure-numbers of the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-components of $\mathbf{s}$.

It is sometimes convenient to regard ∇ as a quasi-vector having $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-components $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$, which in the formation of scalar and vector products behave in the same way as the corresponding components of real vectors.

The divergence operator $\nabla\cdot$. The scalar product of ∇ into the vector point function $\mathbf{s}$ is formed as follows:

$$\begin{aligned}\nabla \cdot \mathbf{s} &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot (s_1\mathbf{i} + s_2\mathbf{j} + s_3\mathbf{k}) \\ &= \frac{\partial s_1}{\partial x} + \frac{\partial s_2}{\partial y} + \frac{\partial s_3}{\partial z}.\end{aligned}$$

Hence:

$$\nabla \cdot \mathbf{s} = \operatorname{div} \mathbf{s}. \tag{3}$$

The symbol $\nabla \cdot$ can therefore be considered as an operator which acting upon a vector point function produces its divergence.

The derivative operator $\mathbf{s} \cdot \nabla$. For the scalar product of $\mathbf{s}$ into ∇ we have:

$$\mathbf{s} \cdot \nabla = (s_1\mathbf{i} + s_2\mathbf{j} + s_3\mathbf{k}) \cdot \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right).$$

Upon expansion of the indicated product on the right we find:

$$\mathbf{s} \cdot \nabla = s_1\frac{\partial}{\partial x} + s_2\frac{\partial}{\partial y} + s_3\frac{\partial}{\partial z}. \tag{4}$$

Hence, $\mathbf{s} \cdot \nabla$ is a new scalar differential operator.

If u is a scalar point function and $\mathbf{v}$ a vector point function, then:

$$(\mathbf{s} \cdot \nabla)\, u = s_1\frac{\partial u}{\partial x} + s_2\frac{\partial u}{\partial y} + s_3\frac{\partial u}{\partial z}; \tag{5}$$

$$(\mathbf{s} \cdot \nabla)\, \mathbf{v} = s_1\frac{\partial \mathbf{v}}{\partial x} + s_2\frac{\partial \mathbf{v}}{\partial y} + s_3\frac{\partial \mathbf{v}}{\partial z}. \tag{6}$$

If $\mathbf{s}$ be a unit vector, then s_1, s_2, s_3 will be the direction cosines of $\mathbf{s}$, and $(\mathbf{s} \cdot \nabla)\, u$ will be equal to the directional derivative of u in the direction of $\mathbf{s}$, viz. $\partial u/\partial s$; and $(\mathbf{s} \cdot \nabla)\, \mathbf{v}$ will be the directional derivative of $\mathbf{v}$ in the direction of $\mathbf{s}$, viz. $\partial \mathbf{v}/\partial s$.

It should be noted that:

$$(\mathbf{s} \cdot \nabla)\, u = \mathbf{s} \cdot (\nabla u). \tag{7}$$

Assuming $\mathbf{s}$ to be a unit vector, this equation states that the directional derivative of u in the direction of $\mathbf{s}$ is equal to the scalar value of the component of the gradient of u in the same direction.

The curl operator $\nabla\times$. In forming the vector product of ∇ into a vector point function $\mathbf{v}$, we find:

$$\begin{aligned}\nabla \times \mathbf{v} &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \times (v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) \\ &= \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}.\end{aligned}$$

Hence:

$$\nabla \times \mathbf{v} = \operatorname{curl} \mathbf{v}. \tag{8}$$

The symbol $\nabla\times$ can, therefore, be regarded as an operator which acting on a vector point function produces its curl.

The Laplacian operator ∇^2. For the scalar product of ∇ into itself we have:

$$\nabla \cdot \nabla = \nabla^2 = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right) \cdot \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right).$$

Hence:

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{9}$$

This is the well known Operator of Laplace.

It acts upon a scalar point function u to give:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}. \tag{10}$$

This is the same result which is obtained when the operator $\nabla \cdot$ acts upon the gradient of u. For:

$$\begin{aligned} \nabla \cdot (\nabla u) &= \text{div grad u} \\ &= \text{div}\left(\mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z}\right), \end{aligned}$$

and hence:

$$\nabla \cdot (\nabla u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \nabla^2 u. \tag{11}$$

If $\nabla^2 u$ vanishes at all points throughout a given region, then u must satisfy the partial differential equation of the second order:

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial u}{\partial z^2} = 0, \tag{12}$$

at all points of the region. This is the celebrated equation of Laplace.

Operating with ∇^2 upon a vector point function $\mathbf{v}$, we find:

$$\begin{aligned} \nabla^2 \mathbf{v} = \quad & \frac{\partial^2 \mathbf{v}}{\partial x^2} + \frac{\partial^2 \mathbf{v}}{\partial y^2} + \frac{\partial^2 \mathbf{v}}{\partial z^2} \\ & \left(\frac{\partial^2 v_1}{\partial x^2} + \frac{\partial^2 v_1}{\partial y^2} + \frac{\partial^2 v_1}{\partial z^2}\right)\mathbf{i} \\ & + \left(\frac{\partial^2 v_2}{\partial x^2} + \frac{\partial^2 v_2}{\partial y^2} + \frac{\partial^2 v_2}{\partial z^2}\right)\mathbf{j} \\ & + \left(\frac{\partial^2 v_3}{\partial x^2} + \frac{\partial^2 v_3}{\partial y^2} + \frac{\partial^2 v_3}{\partial z^2}\right)\mathbf{k}. \end{aligned}$$

Hence:

$$\nabla^2 \mathbf{v} = \mathbf{i}\nabla^2 v_1 + \mathbf{j}\nabla^2 v_2 + \mathbf{k}\nabla^2 v_3. \tag{13}$$

§38

Special Examples Involving Differential Operators

To further exemplify the use of the operator ∇, let us obtain the result of operating with it upon the scalar point function r^n, where r is the magnitude of the position vector $\mathbf{r}$ of a field-point $P(x, y, z)$:

$$\begin{aligned}\nabla r^n &= \mathbf{i}\frac{\partial r^n}{\partial x} + \mathbf{j}\frac{\partial r^n}{\partial y} + \mathbf{k}\frac{\partial r^n}{\partial z}\\ &= \mathbf{i}nr^{n-1}\frac{\partial r}{\partial x} + \mathbf{j}nr^{n-1}\frac{\partial r}{\partial y} + \mathbf{k}nr^{n-1}\frac{\partial r}{\partial z}\\ &= nr^{n-1}\left(\frac{\mathbf{i}x + y\mathbf{j} + z\mathbf{k}}{r}\right).\end{aligned}$$

Hence:

(14) $$\nabla r^n = nr^{n-1}\mathbf{r}_0.$$

Operating with $\nabla \cdot \nabla$ upon r^n, we find:

$$\begin{aligned}\nabla \cdot \nabla r^n = \nabla \cdot (\nabla r^n) &= \text{div } \nabla r^n\\ &= \text{div } (nr^{n-1}\mathbf{r}_0)\\ &= \text{div }\left(nr^{n-1}\frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{r}\right)\\ &= n\left[\frac{\partial}{\partial x}(r^{n-2}x) + \frac{\partial}{\partial y}(r^{n-2}y) + \frac{\partial}{\partial z}(r^{n-2}z)\right.\\ &= n(n-2)\,r^{n-3}\frac{x^2 + y^2 + z^2}{r} + 3nr^{n-2}.\end{aligned}$$

Hence:

(15) $$\nabla^2 r^n = n(n+1)\,r^{n-2}.$$

If $n = -1$, then:

(16) $$\nabla^2 \frac{1}{r} = 0, \qquad r \neq 0.$$

The function $1/r$, therefore, satisfies the equation of Laplace except at the origin.

The following special formulas, in which $\mathbf{a}$ denotes a constant vector, relating to differential operations on the position-vector $\mathbf{r}$, are left for verification to the reader:

$$\nabla \cdot \mathbf{r} = \text{div } \mathbf{r} = 3, \qquad \nabla \times (\mathbf{r} \times \mathbf{a}) = \text{curl } \mathbf{r} \times \mathbf{a} = -2\mathbf{a},$$

$$\nabla \times \mathbf{r} = \text{curl } \mathbf{r} = 0, \qquad \nabla \cdot r\mathbf{a} = \text{div } r\mathbf{a} = \frac{\mathbf{r} \cdot \mathbf{a}}{r},$$

$$\nabla(\mathbf{r} \cdot \mathbf{a}) = \text{grad } \mathbf{r} \cdot \mathbf{a} = \mathbf{a}, \qquad \nabla \times r\mathbf{a} = \text{curl } r\mathbf{a} = \frac{\mathbf{r} \times \mathbf{a}}{r},$$

$$\nabla \cdot (\mathbf{r} \times \mathbf{a}) = \text{div } \mathbf{r} \times \mathbf{a} = 0, \qquad (\mathbf{a} \cdot \nabla)\,\mathbf{r} = \mathbf{a} \cdot (\nabla r) = \mathbf{a} \cdot \text{grad } r = \frac{\mathbf{a} \cdot \mathbf{r}}{r}.$$

§39

Expansion Formulas Involving the Gradients of Scalar Point Functions and the Divergences and the Curls of Vector Point Functions

The following formulas are collected for convenience of reference. Each of them can easily be verified by the method of semi-Cartesian expansion, as will be exemplified in the case of one of them.

Let u and v be scalar point functions, and $\mathbf{u}$ and $\mathbf{v}$ vector point functions. Then:

(1) $\operatorname{grad}(u+v) \equiv \nabla(u+v) = \nabla u + \nabla v$

(2) $\operatorname{div}(\mathbf{u}+\mathbf{v}) \equiv \nabla\cdot(\mathbf{u}+\mathbf{v}) = \nabla\cdot\mathbf{u} + \nabla\cdot\mathbf{v}$

(3) $\operatorname{curl}(\mathbf{u}+\mathbf{v}) \equiv \nabla\times(\mathbf{u}+\mathbf{v}) = \nabla\times\mathbf{u} + \nabla\times\mathbf{v}$

(4) $\operatorname{grad}(uv) \equiv \nabla(uv) = v\nabla u + u\nabla v$

(5) $\operatorname{div}(u\mathbf{v}) \equiv \nabla\cdot(u\mathbf{v}) = \nabla u\cdot\mathbf{v} + u\nabla\cdot\mathbf{v}$

(6) $\operatorname{curl}(u\mathbf{v}) \equiv \nabla\times(u\mathbf{v}) = \nabla u\times\mathbf{v} + u\nabla\times\mathbf{v}$

(7) $\operatorname{grad}(\mathbf{u}\cdot\mathbf{v}) \equiv \nabla(\mathbf{u}\cdot\mathbf{v}) = (\mathbf{u}\cdot\nabla)\mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{u} + \mathbf{u}\times(\nabla\times\mathbf{v}) + \mathbf{v}\times(\nabla\times\mathbf{u})$

(8) $\operatorname{div}(\mathbf{u}\times\mathbf{v}) \equiv \nabla\cdot(\mathbf{u}\times\mathbf{v}) = \mathbf{v}\cdot\nabla\times\mathbf{u} - \mathbf{u}\cdot\nabla\times\mathbf{v}$

(9) $\operatorname{curl}(\mathbf{u}\times\mathbf{v}) \equiv \nabla\times(\mathbf{u}\times\mathbf{v}) = \mathbf{u}(\nabla\cdot\mathbf{v}) - \mathbf{v}(\nabla\cdot\mathbf{u}) + (\mathbf{v}\cdot\nabla)\mathbf{u} - (\mathbf{u}\cdot\nabla)\mathbf{v}$

(10) $\operatorname{curl}^2\mathbf{u} \equiv \nabla\times(\nabla\times\mathbf{u}) = \nabla\nabla\cdot\mathbf{u} - \nabla^2\mathbf{u}$

(11) $\operatorname{div}\operatorname{grad} u \equiv \nabla\cdot\nabla u = \nabla^2 u$

(12) $\operatorname{curl}\operatorname{grad} u \equiv \nabla\times\nabla u = 0$

(13) $\operatorname{div}\operatorname{curl}\mathbf{u} \equiv \nabla\cdot(\nabla\times\mathbf{u}) = 0$

[Special attention is called to formula (12), which asserts that the curl of the gradient of a scalar point function vanishes, and to formula (13), which states that the divergence of the curl of a vector point function vanishes.]

By way of example we give the proof formula (10):

$$\operatorname{curl}^2\mathbf{u} \equiv \nabla\times(\nabla\times\mathbf{u})$$

$$= \nabla\times\left[\left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right)\mathbf{k}\right]$$

$$\left[\frac{\partial}{\partial y}\left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right) - \frac{\partial}{\partial z}\left(\frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}\right)\right]\mathbf{i}$$
$$+ \left[\frac{\partial}{\partial z}\left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}\right) - \frac{\partial}{\partial x}\left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right)\right]\mathbf{j}$$
$$+ \left[\frac{\partial}{\partial x}\left(\frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}\right) - \frac{\partial}{\partial y}\left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}\right)\right]\mathbf{k}$$
$$= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)\left(\frac{\partial u_1}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u_3}{\partial z}\right)$$
$$- \left[\left(\frac{\partial^2 u_1}{\partial x^2} + \frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2 u_1}{\partial z^2}\right)\mathbf{i}\right.$$
$$+ \left(\frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial y^2} + \frac{\partial^2 u_2}{\partial z^2}\right)\mathbf{j}$$
$$+ \left(\frac{\partial^2 u_3}{\partial x^2} + \frac{}{\partial y^2} + \frac{\partial^2 u_3}{\partial z^2}\right)\mathbf{k}.$$

Hence:

$$\text{curl}^2\,\mathbf{u} \equiv \nabla\times(\nabla\times\mathbf{u}) = \nabla\nabla\cdot\mathbf{u} - \nabla^2\mathbf{u},$$

which is formula (10).

EXERCISES ON CHAPTER III

1. Prove the formulas given in the text at the end of §38.

2. If r is the magnitude of the position-vector $\mathbf{r}$ of a field point, show that:

$$\text{div } r^n\mathbf{r} = (n+3)\,r^n;$$
$$\text{curl } r^n\mathbf{r} = 0.$$

3. A scalar point function u is specified by the equation:

$$u = \mathbf{c}\cdot\mathbf{r} + \frac{1}{2}\log_e(\mathbf{c}\times\mathbf{r})^2,$$

where $\mathbf{r}$ is the position vector of a field point and $\mathbf{c}$ is a constant unit vector. Find analytically the gradient of u, and then the divergence and curl of the gradient. Construct a diagram showing the level surfaces of u and the vector lines of the gradient of u.

4. Find the vector equation of the tangent plane at any point on the surface whose equation is $f(\mathbf{r}) = \text{const.}$; also find the equation of the normal to the tangent plane at the point of tangency.

5. If u is a scalar point function, show that:

$$f(u)\,\nabla u = \nabla\int f(u)\,du.$$

6. If:

$$\rho\mathbf{F} = \nabla p,$$

where ρ, p, $\mathbf{F}$ are point functions, show that:

$$\mathbf{F} \cdot \operatorname{curl} \mathbf{F} = 0.$$

7. If $\mathbf{v}$ is the velocity of a point in a rigid body which is rotating with angular velocity $\boldsymbol{\omega}$ about an axis, show that:

$$\boldsymbol{\omega} = \frac{1}{2} \operatorname{curl} \mathbf{v}.$$

8. If $\mathbf{v}$ is a vector point function, show that:

$$(\mathbf{v} \cdot \nabla)\, \mathbf{v} = \frac{1}{2} \nabla(\mathbf{v} \cdot \mathbf{v}) - \mathbf{v} \times \operatorname{curl} \mathbf{v}.$$

9. Prove the formula:

$$\mathbf{A} \times \mathbf{B} \cdot \operatorname{curl} \mathbf{v} = (\mathbf{A} \cdot \nabla)\, \mathbf{V} \cdot \mathbf{B} - (\mathbf{B} \cdot \nabla)\, \mathbf{V} \cdot \mathbf{A},$$

where $\mathbf{v}$ is a vector point function and $\mathbf{A}$ and $\mathbf{B}$ are constant vectors.

10. In passing from one rectangular Cartesian system of axes to another, show that the ∇ operator is an invariant; show also by actual transformation that the forms for the divergence and the curl of a vector point function are invariants.

CHAPTER IV

THEOREMS ON THE TRANSFORMATION OF VOLUME INTEGRALS INTO SURFACE INTEGRALS AND OF SURFACE INTEGRALS INTO LINE INTEGRALS—APPLICATIONS

§40

Derivation of Fundamental Theorems

We consider first the transformation of certain volume integrals taken throughout a region of space V into equivalent surface integrals taken over a surface S bounding V.

Let u and $\mathbf{v}$ respectively denote a scalar and a vector point function, these functions together with their first space derivatives being supposed finite, continuous, and single-valued. We shall show that:

$$\int_V \operatorname{grad} u d\tau = \int_S u\mathbf{n}d\sigma, \tag{1}$$

$$\int_V \operatorname{div} \mathbf{v}d\tau = \int_S \mathbf{n}\cdot\mathbf{v}d\sigma, \tag{2}$$

$$\int_V \operatorname{curl} \mathbf{v}d\tau = \int_S \mathbf{n}\times\mathbf{v}d\sigma, \tag{3}$$

where $\mathbf{n}$ is an outward unit normal to an infinitesimal element of S of magnitude $d\sigma$, and $d\tau$ is the magnitude of an infinitesimal element of V.

We suppose the region V subdivided into infinitesimal elements. Each of the surface integrals must be equal to the sum of corresponding integrals taken over the surfaces (ω) of these elements. For, except for the portions of the surfaces of the elements which abut on S, the bounding surface of each element will be integrated over twice, but with opposite directions for the outward unit normal $(\boldsymbol{\varepsilon})$ in the two integrations, cancellation resulting, while the sum of the contributions of the excepted portions will evidently be equivalent to the original surface integral over S. Consequently:

$$\begin{aligned}
\int_S u\mathbf{n}d\sigma &= \sum\int_\omega u\boldsymbol{\varepsilon}d\sigma,\\
\int_S \mathbf{n}\cdot\mathbf{v}d\sigma &= \sum\int_\omega \boldsymbol{\varepsilon}\cdot\mathbf{v}d\sigma,\\
\int_S \mathbf{n}\times\mathbf{v}d\sigma &= \sum\int_\omega \boldsymbol{\varepsilon}\times\mathbf{v}d\sigma,
\end{aligned} \tag{4}$$

where Σ denotes summation throughout the region V. But from equations (9), (10), and (11) of Art. 36 we have:

$$
\begin{aligned}
&\sum\int_{\omega} u\, d\sigma &&= \sum \operatorname{grad} u d\tau \equiv \int_V \operatorname{grad} u d\tau,\\
(5)\qquad &\sum\int_{\omega} \boldsymbol{\varepsilon}\cdot \mathbf{v} d\sigma &&= \sum \operatorname{div} \mathbf{v} d\tau \equiv \int_V \operatorname{div} \mathbf{v} d\tau,\\
&\sum\int_{\omega} \boldsymbol{\varepsilon}\times \mathbf{v} d\sigma &&= \sum \operatorname{curl} \mathbf{v} d\tau \equiv \int_V \operatorname{curl} \mathbf{v} d\tau.
\end{aligned}
$$

From equations (4) and (5) the validity of equations (1), (2), and (3) follows at once.

Theorem (2) is of particular importance, and is commonly known as the Divergence Theorem or as Gauss's Integral Theorem.

We next consider the transformation of two important surface integrals into equivalent line integrals, the surface integrals being taken over an unclosed surface S, and the line integrals around the contour C bounding S. We shall show that:

$$(6)\qquad \int_S \mathbf{n}\times \operatorname{grad} u d\sigma = \int_C u d\mathbf{r},$$

$$(7)\qquad \int_S \mathbf{n}\cdot \operatorname{curl} \mathbf{v} d\sigma = \int_C \mathbf{v}\cdot d\mathbf{r},$$

where $\mathbf{n}$ specifies a unit normal to an element of S of magnitude $d\sigma$ from the positive side of S, and $d\mathbf{r}$ an element of the contour in the direction around it which is reckoned positive. By convention:

The positive side of an unclosed surface is related to the positive direction around its contour as follows: imagine a right-handed screw, with its axis coincident with a tangent to the contour at any point, to advance in the positive direction of the contour at the point—then the threads of the screw will pierce the surface from its negative to its positive side.

Now suppose the surface S to be subdivided into infinitesimal elements of area. Then the line integral around the contour in equation (6) or (7) is equivalent to the sum of the line integrals taken in the same direction around the contours (c) of all the infinitesimal elements into which S is supposed subdivided. For, excepting those portions of the contours of the elements which abut on the contour C of S, the contours of the elements will be integrated over twice, but in opposite directions in the two cases, so that cancellation will result, save for the excepted portions, which collectively supply an integration equivalent to that around the contour itself. Consequently:

$$\begin{aligned}\int_C u d\mathbf{r} &= \sum \int_c u d\mathbf{s},\\ \int_C \mathbf{v} \cdot d\mathbf{r} &= \sum \int_c \mathbf{v} \cdot d\mathbf{s},\end{aligned} \tag{8}$$

where Σ specifies summation over S, c the elementary contour bounding an element of S, and $d\mathbf{s}$ an element of this contour in the positive direction around it. But from equations (12) and (13), Art. 36, we have:

$$\begin{aligned}\sum \int_c u d\mathbf{s} &= \sum \mathbf{n} \times \operatorname{grad} u d\sigma = \int_S \mathbf{n} \times \operatorname{grad} u d\sigma,\\ \sum \int_c \mathbf{v} \cdot d\mathbf{s} &= \sum \mathbf{n} \cdot \operatorname{curl} \mathbf{v} d\sigma = \int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{v} d\sigma.\end{aligned} \tag{9}$$

From equations (8) and (9) the validity of equations (6) and (7) follows at once.

Theorem (7) is known as Stokes's Theorem, and is of fundamental importance. The converse of this theorem is also true:

If a vector point function $\mathbf{u}$ is related to another vector point function $\mathbf{v}$ in such manner that the surface integral of the normal component of $\mathbf{u}$ over any surface S is equal to the line integral of $\mathbf{v}$ around the contour C bounding the surface, then $\mathbf{u}$ must be the curl of $\mathbf{v}$. For:

$$\begin{aligned}\int_S \mathbf{n} \cdot (\mathbf{u} - \operatorname{curl} \mathbf{v})\, d\sigma &= \int_S \mathbf{n} \cdot \mathbf{u} d\sigma - \int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{v} d\sigma\\ &= \int_C \mathbf{v} \cdot d\mathbf{r} - \int_C \mathbf{v} \cdot d\mathbf{r},\\ &= 0,\end{aligned}$$

and, since this result is true for an arbitrary choice of the surface S, it is necessary that: $\mathbf{u} = \operatorname{curl} \mathbf{v}$.

§41

On the Connectivity of Space

A region of space is said to be Singly-Connected or Acyclic when it is possible through deformation to bring any two lines connecting any two points within it into coincidence without either line cutting out of the region. Any region for which this is not possible is said to be Multiply-Connected or Cyclic.

The region within an anchor ring (see Fig. 31) is an example of a cyclic region (doubly-connected). In this case the region

can be converted into an acyclic (singly-connected) region by the device of assuming a section of the ring (indicated by the dotted line) to constitute a barrier across which no line can pass; the two sides of the barrier being then considered as parts of the surface bounding the acyclic region thus produced.

In like manner a doubly-connected area can be reduced by a line barrier to a singly-connected area.

By the device of introducing one or more barriers, any cyclic region can be reduced to an acyclic region.

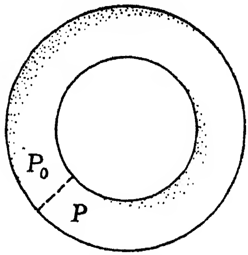


Fig. 31.

Suppose, now, that the anchor ring is in the field of a vector point function $\mathbf{v}$, and that:

$$\text{curl } \mathbf{v} = 0, \text{ everywhere inside the ring,}$$
$$\text{curl } \mathbf{v} \neq 0, \text{ everywhere outside the ring.}$$

Let C denote a closed path of integration inside the ring. Then, by Stokes's theorem:

$$\int_C \mathbf{v} \cdot d\mathbf{r} = \int_S \mathbf{n} \cdot \text{curl } \mathbf{v} d\sigma$$

Unless the region inside the ring be made acyclic through the introduction of a barrier, it will be possible to draw the path C so as to encircle the hole of the ring, in which case the surface S must lie, in part at any rate, in the region outside the ring in which **curl v** does not vanish everywhere; and the surface integral will therefore have in general a value different from zero. If we call this value A, then:

$$\int_C \mathbf{v} \cdot d\mathbf{r} = A, \tag{1}$$

if the path C encircle the hole of the ring; if the path C does not do this, the integral will have a zero value, since it will then be possible to draw the surface S so as to lie wholly within the region inside the ring for which curl $\mathbf{v}$ is supposed zero. It follows that, if the region within the ring be made acyclic through the introduction of the barrier, then, however the path C be drawn:

$$\int \mathbf{v} \cdot d\mathbf{r} = 0. \tag{2}$$

From this result it follows that the integrals of $\mathbf{v} \cdot d\mathbf{r}$ along *any* two lines within the ring connecting two points P_0 and P have a

common value, and hence that we can define a single-valued scalar point function by the equation:

$$\phi = \int_{P_0}^{P} \mathbf{v} \cdot d\mathbf{r}, \tag{3}$$

where P_0 is a fixed point and P any other point within the ring.

If the region within the ring be not reduced to an acyclic region, the function ϕ must, in virtue of equation (1), be replaced by a multiple-valued function with values at P which differ by integer multiples of A (which can be shown to have a constant value), depending upon the number of times which the path of integration from P_0 to P encircles the hole. The proof of these statements is left as an exercise to the reader.

A physical example of a multiple-valued scalar point function of this sort is furnished by the magnetic potential due to a distribution of electric currents, provided the region about the currents is not made acyclic through the introduction of appropriate barriers.

We shall assume in general that the regions with which we deal in the future are acyclic to begin with or have been made so by the introduction of appropriate barriers.

§42

Discontinuous Scalar and Vector Point Functions

The functions which we have dealt with in deriving the preceding general transformation theorems for volume, surface, and line integrals have been supposed, together with their first space derivatives, to be finite, single-valued, and continuous, and in using these theorems it is therefore necessary to pay special attention to any discontinuities possessed by functions to which it is proposed to apply them. It will suffice to illustrate the matter by a single example.

In the case of the divergence theorem:

$$\int_V \operatorname{div} \mathbf{v} d\tau = \int_S \mathbf{n} \cdot \mathbf{v} d\sigma,$$

suppose that $\mathbf{v} = \mathbf{r}/r^3$, where $\mathbf{r}$ is the position-vector with respect to a fixed point O of any point of the region V or the surface S, the point O being supposed within V. The function $\mathbf{v}$ will then take on infinite values at the point O, for which $r = 0$. Hence, before using the theorem, the point O, which gives trouble, must be

excluded from the field of integration by surrounding it by a small closed surface which is then to be considered as part of the surface bounding V. A case of this sort will be treated in detail in the following article.

Multiple-valued functions must be treated with the same caution as discontinuous functions.

§43
Gauss's Theorem

From the general transformation theorems derived in Art. 40 special theorems of great importance can be found. The first of these which we shall consider is known as Gauss's Theorem.

This well known theorem can be derived from the divergence theorem, equation (2), Art. 40, viz.:

$$\int_V \operatorname{div} \mathbf{v} d\tau = \int_S \mathbf{n} \cdot \mathbf{v} d\sigma.$$

We take:

$$\mathbf{v} = -\nabla \frac{1}{r} = \frac{\mathbf{r}_0}{r^2} = \frac{\mathbf{r}}{r^3}, \tag{1}$$

where r is the magnitude of the position-vector $\mathbf{r}$ with respect to an arbitrary origin O of a field-point P; whereupon we obtain the formula:

$$\int_S \frac{\mathbf{n} \cdot \mathbf{r}_0}{r^2} d\sigma = \int_V \operatorname{div} \frac{\mathbf{r}}{r^3} d\tau. \tag{2}$$

This formula is valid provided O lies without the surface S bounding the region V, the restriction regarding validity arising from the fact that all theorems in the present chapter have been derived on the assumption that the point functions considered, together with their first space derivatives, are continuous.

If, then, the point O is outside S:

$$\operatorname{div} \frac{\mathbf{r}}{r^3} = \frac{1}{r^3} \operatorname{div} \mathbf{r} + \nabla \frac{1}{r^3} \cdot \mathbf{r} = \frac{3}{r^3} - \frac{3}{r^3} = 0,$$

and the formula therefore gives:

$$\int_S \frac{\mathbf{n} \cdot \mathbf{r}_0}{r^2} d\sigma = 0, \text{ for } O \text{ outside } S. \tag{3}$$

If O be a point within S, then at this point $1/r$ will have an infinite value. But if we suppose the point O to be the center of a

small sphere with radius ϵ and surface S', formula (3) will be valid for the region bounded by the original surface S and the surface S' of the small sphere. Consequently:

$$\int_S \frac{\mathbf{n}\cdot\mathbf{r}_0}{r^2}\,d\sigma + \int_{S'} \frac{\mathbf{n}\cdot\mathbf{r}_0}{r^2}\,d\sigma' = 0,$$

but:

$$\int_{S'} \frac{\mathbf{n}\cdot\,{}_0}{r^2}\,d\sigma' = \frac{\mathbf{n}\cdot\mathbf{r}_0}{\epsilon^2}\int_{S'} d\sigma' = -\frac{4\pi\epsilon^2}{\epsilon^2} = -4\pi,$$

and hence:

(4) $$\int_S \frac{\mathbf{n}\cdot\mathbf{r}_0}{r^2}\,d\sigma = 4\pi, \text{ for } O \text{ inside } S.$$

Formulas (3) and (4) constitute the statement of a theorem often called Gauss's Theorem.

§44
Green's Theorem

This celebrated theorem is easily derived with the aid of the divergence theorem:

$$\int_V \operatorname{div} \mathbf{v}\,d\tau = \int_S \mathbf{n}\cdot\mathbf{v}\,d\sigma.$$

Let $\mathbf{v} = W\mathbf{w}$, where W is a scalar and $\mathbf{w}$ a vector point function.[1)] Then, by formula (5), Art. 39:

$$\operatorname{div} \mathbf{v} = \operatorname{div} W\mathbf{w} = W \operatorname{div} \mathbf{w} + \nabla W\cdot\mathbf{w}.$$

Hence, by the divergence theorem:

$$\int_V W \operatorname{div} \mathbf{w}\,d\tau + \int_V \nabla W\cdot\mathbf{w}\,d\tau = \int_S W\mathbf{n}\cdot\mathbf{w}\,d\sigma.$$

Now suppose $\mathbf{w}$ to be the gradient of a scalar point function U, so that $\mathbf{w} = \nabla U$. Then, by formula (11), Art. 39:

$$\operatorname{div} \mathbf{w} = \operatorname{div} \nabla U = \nabla^2 U,$$

and the preceding equation therefore gives:

(1) $$\int_V W\nabla^2 U\,d\tau + \int_V \nabla U\cdot\nabla W\,d\tau = \int_S W\mathbf{n}\cdot\nabla U\,d\sigma.$$

1) These functions, together with the function U introduced below, and their first space derivatives are supposed finite, continuous, and single-valued within the region V.

On account of the symmetry of the second integral in U and W these quantities can be interchanged in each of the three integrals without prejudice to the equality sign. Consequently:

$$(2) \qquad \int_V U\nabla^2 W d\tau + \int_V \nabla W \cdot \nabla U d\tau = \int_S U\mathbf{n} \cdot \nabla W d\sigma.$$

From equations (1) and (2), by subtraction, we get:

$$(3) \qquad \int_V (U\nabla^2 W - W\nabla^2 U)\, d\tau = \int_S \mathbf{n} \cdot (U\nabla W - W\nabla U)\, d\sigma.$$

Formula (1) is known as Green's Theorem in the First Form, and formula (3) as Green's Theorem in the Second Form.

§45

Solenoidal Vector Point Functions

If throughout a given region a vector point function $\mathbf{v}$ satisfy the condition:

$$(1) \qquad \operatorname{div} \mathbf{v} = 0,$$

the function $\mathbf{v}$ is said to have a solenoidal distribution within the region, and is called a Solenoidal Vector Point Function.

If S denote any closed surface in the region, then, by the divergence theorem:

$$(2) \qquad \int_S \mathbf{n} \cdot \mathbf{v} d\sigma = 0,$$

where $\mathbf{n}$ denotes a unit vector in the direction of an outward drawn normal to S, and $d\sigma$ the magnitude of an infinitesimal element of S.

If S be in the form of a tube (a Vector Tube) whose generators are vector lines of $\mathbf{v}$, with ends S_1 and S_2, the lateral walls will contribute nothing to the integral, since for them $\mathbf{n} \cdot \mathbf{v} = 0$. The contribution of the two ends is expressed as follows:

$$\int_{S_1} \mathbf{n} \cdot \mathbf{v} d\sigma_1 + \int_{S_2} \mathbf{n} \cdot \mathbf{v} d\sigma_2 = 0.$$

Now suppose the vector lines of $\mathbf{v}$ to be directed from S_1 towards S_2, and let $\mathbf{n}_1$ specify an *inward* drawn unit normal at S_1 and $\mathbf{n}_2$ an *outward* drawn unit normal at S_2; then, by the last equation:

$$(3) \qquad \int_{S_1} \mathbf{n}_1 \cdot \mathbf{v} d\sigma_1 = \int_{S_2} \mathbf{n}_2 \cdot \mathbf{v} d\sigma_2.$$

If we define the flux of a vector $\mathbf{v}$ through an unclosed surface S as follows:

$$\text{(4)} \qquad \text{Flux of } \mathbf{v} \text{ through } S = \int_S \mathbf{n} \cdot \mathbf{v} d\sigma,$$

then equation (3) states that the flux of the vector through S_1 is equal to the flux of the vector through S_2. In other words, the flux of a solenoidal vector across a section of a corresponding vector tube is constant along the tube.

If the flux through a vector tube be equal to unity, it is called a Unit Vector Tube. Evidently, the number of unit tubes of a vector $\mathbf{v}$ entering a closed surface S must be equal to the number leaving S, provided condition (1) is satisfied by $\mathbf{v}$ throughout the region bounded by S.

Examples of solenoidal vector point functions are furnished by: the velocity of an incompressible fluid, the gravitational field-intensity in a region containing no gravitating matter, and the electric field-intensity in a region containing no electric charges.

§46
Lamellar Vector Point Functions

If throughout a given region a vector point function $\mathbf{v}$ satisfy the condition:

$$\text{(1)} \qquad \operatorname{curl} \mathbf{v} = 0,$$

the function $\mathbf{v}$ is said to have a lamellar distribution within the region, and is called a Lamellar Vector Point Function.

When this condition is satisfied in an acyclic region and only when, there must exist a single-valued scalar point function, ϕ say, such that:[1)]

$$\text{(2)} \qquad \mathbf{v} = \nabla\phi,$$

as can be proved as follows:

By Stoke's Theorem, equation (7), Art. 40:

$$\int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{v} d\sigma = \int_C \mathbf{v} \cdot d\mathbf{r}.$$

Hence, for all possible closed contours in any acyclic region:

$$\int_C \mathbf{v} \cdot d\mathbf{r} = 0,$$

if and only if $\operatorname{curl} \mathbf{v} = 0$ throughout the region. In this case, if P_0 denote a fixed point and $P(x, y, z)$ any other point in the given

[1)] If the region is not acyclic a function ϕ such that $\mathbf{v} = \nabla\phi$ will still exist, if $\operatorname{curl} \mathbf{v} = 0$ throughout the region, but it will not in general be single-valued.

region, and if (*a*) and (*b*) denote any two paths within the region connecting P_0 and P, then:

$$\int_{(a)} \mathbf{v} \cdot d\mathbf{r} = \int_{(b)} \mathbf{v} \cdot d\mathbf{r}.$$

This equation shows that the value of the integral of $\mathbf{v} \cdot d\mathbf{r}$ along an acyclic path from P_0 to P is independent of the path, or, in other words, that the value of the integral depends simply upon the co-ordinates x, y, z of P and the co-ordinates of P_0, the latter point being supposed fixed. We can therefore write:

$$\int_{P_0}^{P} \mathbf{v} \cdot d\mathbf{r} = \phi(x, y, z), \tag{3}$$

where ϕ is a scalar point function. Hence, at any point $P(x, y, z)$ in the given region:

$$d\phi = \mathbf{v} \cdot d\mathbf{r}. \tag{4}$$

But we have also, by equation (6), Art. 33:

$$d\phi = \nabla\phi \cdot d\mathbf{r}.$$

Consequently:

$$\mathbf{v} \cdot d\mathbf{r} = \nabla\phi \cdot d\mathbf{r}. \tag{5}$$

This equation is valid for all possible values of $d\mathbf{r}$. Hence, equation (2) must be valid.

It should be further remarked that curl $\mathbf{v} = 0$ constitutes the necessary and sufficient condition that $v_1dx + v_2dy + v_3dz$ shall be a perfect differential of some function, ϕ say, of the co-ordinates x, y, z, since the vector condition curl $\mathbf{v} = 0$ is equivalent to the familiar Cartesian conditions for the existence of ϕ:

$$\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z} = 0, \qquad \frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x} = 0, \qquad \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} = 0.$$

Lamellar vector point functions are frequently encountered in theoretical physics. Examples are furnished by: the velocity of a fluid in which no vortex motion is present, and by the gravitational, or electric field-intensity, due to a distribution of gravitational matter, or of electric charges.

§47

An Application of Green's Theorem

With the aid of Green's Theorem we shall now show that a vector point function $\mathbf{v}$ is uniquely determined within a region V bounded

by a surface S when its divergence and curl are given throughout V together with its normal component over S.

Let $\mathbf{v}'$ be a second vector point function having the same values as $\mathbf{v}$ for its divergence and curl throughout V and for its normal component over S. Denoting the difference of $\mathbf{v}'$ and $\mathbf{v}$ by $\mathbf{q}$, we shall then have:

$$\operatorname{div}\mathbf{q} = 0, \qquad \operatorname{curl}\mathbf{q} = 0, \text{ throughout } V, \qquad \mathbf{n}\cdot\mathbf{q} = 0, \text{ over } S,$$

where $\mathbf{n}$ is a unit vector in the direction of an outward drawn normal to S.

Since $\operatorname{curl}\mathbf{q} = 0$, therefore:

$$\mathbf{q} = \nabla u,$$

where u is some scalar point function. Since $\operatorname{div}\mathbf{q} = 0$, it follows that:

$$\nabla^2 u = 0, \text{ throughout } V,$$

and, since $\mathbf{n}\cdot\mathbf{q} = 0$, that:

$$\mathbf{n}\cdot\nabla u = 0, \text{ over } S.$$

Now, in the first form of Green's Theorem, equation (1), Art. 44, take $U = W = u$, then, with the aid of the last two equations, we find:

$$\int_V (\nabla u)^2\, d\tau = 0.$$

Hence, since the co-factor of $d\tau$ in the integrand cannot be negative, it follows that:

$$\nabla u = 0, \text{ and therefore } \mathbf{q} = 0,$$

for all points throughout V and over S. Therefore, $\mathbf{v}'$ must equal $\mathbf{v}$, and the uniqueness of $\mathbf{v}$ is established.

§48[1)]

The Hydrodynamical Equation of Continuity

If $\mathbf{q}$ represent the velocity of a moving fluid at any point and ρ its density, then $\rho\mathbf{q}$ will represent its momentum per unit volume at the point. Consider a closed surface S bounding a region V, within the fluid but fixed in space. At any instant the mass of fluid flowing outward per unit of time through an element of surface

1) The present and following article are concerned with physical applications; and the symbols introduced to represent physical quantities are supposed to carry the physical dimensions of these quantities.

dS is equal to that of the fluid which would be contained in an oblique cylinder on the base dS with axis in the direction of $\mathbf{q}$, of slant height q, and of density equal to that existing at the element at the instant under consideration. The mass in question will, therefore, be represented by $\mathbf{n} \cdot \rho\mathbf{q}dS$, where $\mathbf{n}$ is a unit vector in the direction of an outward normal at the element. Hence, the mass of fluid flowing outward per unit time through the closed surface will be represented by

$$\int_S \mathbf{n} \cdot \rho\mathbf{q}dS;$$

and by the divergence theorem:

$$(1) \qquad \int_S \mathbf{n} \cdot \rho\mathbf{q}dS = \int_V \operatorname{div} \rho\mathbf{q}dV.$$

But, admitting that fluid cannot be created or destroyed, the mass in question must also be represented by

$$-\frac{\partial}{\partial t}\int_V \rho dV = -\int_V \frac{\partial \rho}{\partial t} dV,$$

where t denotes the time. Consequently:

$$-\int_V \frac{\partial \rho}{\partial t} dV = \int_V \operatorname{div} \rho\mathbf{q}dV.$$

The closed surface S is arbitrary as regards size, and therefore the volume V which it encloses is also arbitrary in size, and it follows from the last equation that at each point of the fluid:

$$(2) \qquad -\frac{\partial \rho}{\partial t} = \operatorname{div} \rho\mathbf{q}.$$

This equation is valid whether or not the fluid is incompressible. It is a consequence of the assumption of the conservation of mass of the fluid, and is therefore called the Equation of Continuity.

If the fluid be incompressible, then:

$$(3) \qquad \frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x}\frac{dx}{dt} + \frac{\partial \rho}{\partial y}\frac{dy}{dt} + \frac{\partial \rho}{\partial z}\frac{dz}{dt} = \frac{\partial \rho}{\partial t} + \nabla\rho \cdot \mathbf{q} = 0,$$

where $d\rho/dt$ represents the time rate of change in the density of a particle of the moving fluid whose co-ordinates are x, y, z. From equations (2) and (3) it follows that for an incompressible fluid:

$$(4) \qquad \operatorname{div} \mathbf{q} = 0.$$

If, in addition to being incompressible, the fluid is of uniform density throughout, ρ will be independent of the space co-ordinates as well as the time, and, by equations (1) and (4), in this case:

(5) $$\int_S \mathbf{n} \cdot \rho \mathbf{q} dS = 0.$$

This equation expresses that the mass of fluid leaving the closed surface S per unit time is equal to that entering it.

Maxwell's Electromagnetic Field Equations for Free Space

An interesting application of Stokes's Theorem is found in the derivation of two of Maxwell's electromagnetic field equations for free space. These equations, by means of which Maxwell described variable electromagnetic phenomena in a region devoid of ponderable matter, rest ultimately of course upon the results of experiment. They express the behavior in space and time of two fundamental vector quantities **E** and **H** representing respectively the electric force or field-intensity, and the magnetic force or field-intensity.

In the first place, experimental results, by arguments which it is not our purpose to consider here, lead to the inference that in free space both **E** and **H** are to be considered as solenoidal vector point functions, so that:

(1) $$\operatorname{div} \mathbf{E} = 0,$$

(2) $$\operatorname{div} \mathbf{H} = 0.$$

Experiments on electromagnetic induction by Faraday and by Henry indicate the truth of the following statement:

The work which would be done by the forces of an electromagnetic field in free space upon a unit positive charge, if made to pass in the positive direction once around a closed path (C), is proportional to the time rate of decrease of the flux of magnetic force in the positive direction through any surface (S) bounded by the path.[1]

To put this statement in mathematical form, take $d\mathbf{r}$ to represent in magnitude and direction an element of the path in the direction around it reckoned as positive, and **n** as a unit vector in the direction of a normal to the surface from its positive side;[1] the statement can then be expressed by the equation:

(a) $$\int_C \mathbf{E} \cdot d\mathbf{r} = -A \frac{\partial}{\partial t} \int_S \mathbf{n} \cdot \mathbf{H} dS,$$

where A is a positive constant. The integral on the left expresses the electromotive force around the closed path C.

[1] See the convention adopted in Art. 40.

Maxwell, guided by the idea that all electric currents must be closed currents, possessing, therefore, the solenoidal property, and by the results of the electromagnetic experiments of Ampere, was led to make a corresponding statement in which, except for a matter of sign, the electric and magnetic field-intensities exchange rôles. The statement is equivalent to the following:

The work which would be done by the forces of an electromagnetic field in free space upon a unit positive magnetic pole, if made to pass in the positive direction once around the closed path C, is proportional to the time rate of increase of the flux of electric force in the positive direction through the surface S bounded by the path.

The mathematical expression of this statement is as follows:

$$\int_C \mathbf{H} \cdot d\mathbf{r} = B\frac{\partial}{\partial t}\int_S \mathbf{n} \cdot \mathbf{E}dS, \tag{b}$$

where B is a positive constant. The time rate of change of the integral on the right represents a quantity proportional to that called by Maxwell the Displacement Current through the surface S.

The application of Stokes's Theorem which it is proposed to make consists in deriving the two field equations of Maxwell which follow from the statements expressed by (a) and (b).

By Stokes's theorem, equation (7), Art. 40, we have:

$$\int_C \mathbf{E} \cdot d\mathbf{r} = \int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{E}dS,$$

and hence, with the aid of (a):

$$\int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{E}dS = -A\frac{\partial}{\partial t}\int_S \mathbf{n} \cdot \mathbf{H}dS = -A\int_S \mathbf{n} \cdot \frac{\partial \mathbf{H}}{\partial t}\, dS.$$

Since the closed path C, and therefore the surface S, can be chosen in an arbitrary manner, it follows directly from this equation that:

$$\operatorname{curl} \mathbf{E} = -A\frac{\partial \mathbf{H}}{\partial t}.$$

Again, by Stokes's theorem:

$$\int_C \mathbf{H} \cdot d\mathbf{r} = \int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{H}dS;$$

and hence, with the aid of (b):

$$\int_S \mathbf{n} \cdot \operatorname{curl} \mathbf{H}dS = B\frac{\partial}{\partial t}\int_S \mathbf{n} \cdot \mathbf{E}dS = B\int_S \mathbf{n} \cdot \frac{\partial \mathbf{E}}{\partial t}\, dS.$$

As before, since S can be chosen in an arbitrary manner, it follows that:

$$\operatorname{curl} \mathbf{H} = B\frac{\partial \mathbf{E}}{\partial t}.$$

On the so-called rational system of electromagnetic units $A = B = 1/c$, where c is numerically equal to the ratio of a $C.\,G.\,S.$ electromagnetic unit to a $C.\,G.\,S.$ electrostatic unit of charge, and has a numerical value 3×10^{10}, which is that of the velocity of light on the $C.\,G.\,S.$ system of units.

The two field equations of Maxwell which follow from the assumptions expressed by (a) and (b) are, therefore:

(3) $$\operatorname{curl} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t},$$

(4) $$\operatorname{curl} \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}.$$

The equations (1), (2), (3), and (4) really represent a state of propagation with velocity c in free space of $\mathbf{E}$ and $\mathbf{H}$ in accordance with equations of propagation which are easily found as follows:

By taking the curl of both members of equation (3), we get:

$$\operatorname{curl}^2 \mathbf{E} = \operatorname{curl}\left(-\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}\right) = -\frac{1}{c}\frac{\partial}{\partial t}\operatorname{curl} \mathbf{H};$$

and hence, with the aid of equation (4):

$$\operatorname{curl}^2 \mathbf{E} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

But, by formula (10), Art. 39:

$$\operatorname{curl}^2 \mathbf{E} = \nabla \operatorname{div} \mathbf{E} - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E},$$

since, by equation (1), $\operatorname{div} \mathbf{E} = 0$.

Hence:

(5) $$\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} \quad 0.$$

This is the equation of propagation in free space for the electric force $\mathbf{E}$.

In a similar way the corresponding equation for the magnetic force $\mathbf{H}$ can be found to be as follows:

(6) $$\nabla^2 \mathbf{H} - \frac{1}{c^2}\frac{\partial^2 \mathbf{H}}{\partial t^2} \quad 0.$$

EXERCISES ON CHAPTER IV

1. Show that:

$$\frac{1}{3}\int_S \mathbf{n}\cdot\mathbf{r}d\sigma = |V|,$$

where $\mathbf{r}$ is the position-vector of a generic point on any closed surface S bounding a volume V, and $\mathbf{n}$ is a unit vector in the direction of an outward normal to the surface.

2. Show that:

$$\int_c \mathbf{r}\cdot d\mathbf{r} = 0,$$

where $\mathbf{r}$ is the position vector of a generic point on any closed contour c.

3. If u and v are scalar point functions, show that:

$$\int_c u\nabla v\cdot d\mathbf{r} = -\int_c v\nabla u\cdot d\mathbf{r}.$$

4. If u is a scalar and $\mathbf{v}$ a vector point function, show that:

$$\int_S \mathbf{n}\cdot\nabla u\times\mathbf{v}d\sigma = \int_c u\mathbf{v}\cdot d\mathbf{r} - \int_S u\mathbf{n}\cdot\operatorname{curl}\mathbf{v}\,d\sigma,$$

where c denotes the contour bounding the surface S.

5. If u is a scalar and $\mathbf{v}$ a vector point function, show that:

$$\int_V u\operatorname{div}\mathbf{v}d\tau = \int u\mathbf{n}\cdot\mathbf{v}d\sigma - \int_V \nabla u\cdot\mathbf{v}d\tau,$$

where S denotes the surface bounding the volume V.

6. If u and ρ are scalar point functions, and if $\mathbf{v} = \nabla u$ and $\nabla^2 u = -4\pi\rho$, prove the formula:

$$\int_S \mathbf{n}\cdot\mathbf{v}d\sigma = -4\pi\int_V \rho d\tau,$$

where S denotes the surface bounding the volume V.

7. A scalar point function u satisfies the equation:

$$\nabla^2 u + k^2 u = 0, \qquad (k = \text{const.}),$$

and u_0 is the value of the function at the point O. If $\mathbf{r}$ is the position vector with respect to the point O of a generic point, S a closed surface, and $\mathbf{n}$ a unit vector in the outward normal direction to the surface, show that:

$$\int_S\left[\frac{e^{ikr}}{r}\mathbf{n}\cdot\nabla u - u\mathbf{n}\cdot\nabla\left(\frac{e^{ikr}}{r}\right)\right]d\sigma = 4\pi u_{r=0} \text{ or } 0, \qquad (i = \sqrt{-1}),$$

according as the point O is inside or outside the surface S.

8. Prove the following formula, known as Kelvin's generalization of Green's theorem:

$$\int_V \Phi \nabla U \cdot \nabla W d\tau = \int_S U\Phi \mathbf{n} \cdot \nabla W d\sigma - \int_V U \operatorname{div} (\Phi \nabla W)\, d\tau$$
$$= \int_S W\Phi \mathbf{n} \cdot \nabla U d\sigma - \int_V W \operatorname{div} (\Phi \nabla U)\, d\tau,$$

where Φ, U, W are scalar point functions.

9. If $\mathbf{r}$ is the position vector with respect to a point O of a generic point in a volume V bounded by a surface S and U a scalar point function, prove, with the aid of Green's theorem in the second form, the following theorems, known as Green's Formulas:

$$\int_S \mathbf{n} \cdot \left(\frac{1}{r} \nabla U - U \nabla \frac{1}{r}\right) d\sigma - \int_V \frac{\nabla^2 U}{r}\, d\tau = 4\pi U_0 \quad \text{or} \quad 0,$$

according as the point O is inside or outside S, the symbol U_0 representing the value of the function U at the point O.

10. Prove that the necessary and sufficient condition that the surface integral

$$\int_S \mathbf{n} \cdot \mathbf{v} d\sigma$$

shall be capable of transformation into a line integral around the contour bounding the surface S is:

$$\operatorname{div} \mathbf{v} = 0,$$

at all points of the surface.

CHAPTER V

SCALAR AND VECTOR POTENTIAL FUNCTIONS

§50
Definitions of Potential Functions

Let u denote a scalar and $\mathbf{g}$ a vector point function, each subject to certain restrictions which will be specified below. We define another scalar (U) and another vector point function ($\mathbf{G}$) by the equations:

$$(1)\qquad U(x,y,z) = \int \frac{u}{r}\,d\tau, \qquad\qquad (2)\qquad \mathbf{G}(x,y,z) \quad \int \frac{\mathbf{g}}{r}\,d\tau,$$

where r is the magnitude of the position-vector $\mathbf{r}$ of the field-point $P(x, y, z)$ with respect to the Source-point $\Pi(\xi, \eta, \zeta)$ at which the volume element of magnitude $d\tau$ is supposed taken, and where the integrals are supposed taken over all regions of space within which u and $\mathbf{g}$ have values different from zero or, equivalently, over all space. The position-vectors of the points P and Π with respect to an arbitrary origin O are denoted by $\mathbf{s}$ and $\boldsymbol{\varrho}$ respectively. See Fig. 32. The functions u and $\mathbf{g}$ together with their first space derivatives are supposed finite, single-valued, and continuous; furthermore, at infinity u and $\mathbf{g}$ are supposed to vanish to the order $1/r^3$. These conditions are sufficient to ensure the convergence of the integrals defining the functions U and $\mathbf{G}$, and also the differentiability of these functions and of their first space derivatives at all points. For the proof of these statements the reader is referred to any standard treatise[1] dealing with the theory of potential functions.

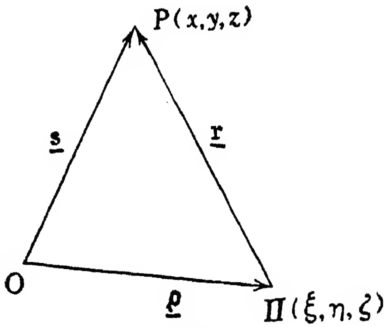


Fig. 32.

At first sight it might appear that the value of U, or $\mathbf{G}$, at the point Π (for which $r = 0$) would be infinite in general, but this is

[1] For example, Gibbs-Wilson, Vector Analysis, p. 208 et seq.

not so. For, using spherical co-ordinates (r, θ, ϕ) with origin at Π, we can write:

$$U = \int\int\int ur \sin\theta\, dr d\theta d\phi; \quad \mathbf{G} = \int\int\int \mathbf{g} r \sin\theta\, dr d\theta d\phi;$$

and these integrals are finite for all points, Π included.

The scalar point function U is called the Scalar Potential of u, and the vector point function $\mathbf{G}$ is called the Vector Potential of $\mathbf{g}$.

It should be carefully noted that in the integrands of the integrals defining U and $\mathbf{G}$ the quantity r is the only one depending upon the position of the field-point $P(x, y, z)$, the other quantities depending only upon the position of the source-point $\Pi(\xi, \eta, \zeta)$.

§51

The Equations of Poisson and of Laplace

For the gradient of U at any field point $P(x, y, z)$ we have:

$$\text{(1)} \qquad \nabla U = \nabla\int\frac{u}{r}\, d\tau = \int\left(\nabla\frac{1}{r}\right) u d\tau = -\int\frac{\mathbf{r}_0}{r^2}\, u d\tau.$$

Hence the contribution to the gradient of U at any field-point $P(x, y, z)$, due to an element of volume of magnitude $d\tau$ at any source-point $\Pi(\xi, \eta, \zeta)$, will be:

$$-\frac{\mathbf{r}_0}{r^2}\, u d\tau.$$

Consider now the surface integral of the normal component of ∇U over any closed surface S bounding a region V. For the contribution to this integral arising from the volume element of magnitude $d\tau$ at $\Pi(\xi, \eta, \zeta)$ we have:

$$-u d\tau \int_S \frac{\mathbf{n} \cdot \mathbf{r}_0}{r^2}\, d\sigma,$$

where $\mathbf{n}$ specifies a unit outward drawn normal to S, and $d\sigma$ the area of an infinitesimal element of S; this, by Gauss's theorem, must be equal to $-4\pi u d\tau$, if Π be inside, and equal to zero, if Π be outside S. Hence, we have for the surface integral in question:

$$\int_S \mathbf{n} \cdot \nabla U d\sigma = -4\pi \int_V u d\tau.$$

But, by the divergence theorem:

$$\int_S \mathbf{n} \cdot \nabla U d\sigma = \int_V \operatorname{div} \nabla U d\tau = \int_V \nabla^2 U d\tau.$$

Since the region V can be chosen arbitrarily, it follows that at any field point $P(x, y, z)$:

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = -4\pi u, \tag{2}$$

where u is evaluated at P, and is therefore to be considered now as a function of x, y, z. This is the celebrated partial differential equation of Poisson. In regions within which $u = 0$ everywhere it degenerates into Laplace's equation:

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0. \tag{3}$$

It has thus been shown that U, the scalar potential of u, must satisfy Poisson's equation at all points, and Laplace's equation at all points for which $u = 0$.

For the vector potential $\mathbf{G}$ of $\mathbf{g}$ and for $\mathbf{g}$ itself we can write:

$$\mathbf{G} = G_1\mathbf{i} + G_2\mathbf{j} + G_3\mathbf{k}, \qquad \mathbf{g} = g_1\mathbf{i} + g_2\mathbf{j} + g_3\mathbf{k},$$

where G_1, G_2, G_3 and g_1, g_2, g_3 are respectively the measure-numbers of the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-components of $\mathbf{G}$ and $\mathbf{g}$; and, from the defining equation for $\mathbf{G}$, we then obtain:

$$G_1 = \int \frac{g_1}{r}\,d\tau, \qquad G_2 = \int \frac{g_2}{r}\,d\tau, \qquad G_3 = \int \frac{g_3}{r}\,d\tau,$$

showing that G_1, G_2, G_3 are the scalar potentials of g_1, g_2, g_3, respectively. Hence, by equation (2):

$$\nabla^2 G_1 = -4\pi g_1, \qquad \nabla^2 G_2 = -4\pi g_2, \qquad \nabla^2 G_3 = -4\pi g_3.$$

Multiplying these equations respectively by $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, and adding, we find that at any field point $P(x, y, z)$ $\mathbf{G}$ must satisfy the equation:

$$\nabla^2 \mathbf{G} = \frac{\partial^2 \mathbf{G}}{\partial x^2} + \frac{\partial^2 \mathbf{G}}{\partial y^2} + \frac{\partial^2 \mathbf{G}}{\partial z^2} = -4\pi \mathbf{g}, \tag{4}$$

where $\mathbf{g}$ is evaluated at P, and is therefore to be considered now as a function of x, y, z. This vector partial differential equation, which is satisfied by the vector potential of $\mathbf{g}$ at all points, is analogous to Poisson's equation. At points for which $\mathbf{g} = 0$ the equation degenerates into the vector partial differential equation:

$$\nabla^2 \mathbf{G} = \frac{\partial^2 \mathbf{G}}{\partial x^2} + \frac{\partial^2 \mathbf{G}}{\partial y^2} + \frac{\partial^2 \mathbf{G}}{\partial z^2} = 0, \tag{5}$$

which is analogous to Laplace's equation for the scalar potential.

From the defining equations (1) and (2), Art. 50, and from equations (2) and (4) above we find:

(6) $$U = -\frac{1}{4\pi}\int \frac{\nabla^2 U}{r}\, d\tau,$$

(7) $$\mathbf{G} = -\frac{1}{4\pi}\int \frac{\nabla^2 \mathbf{G}}{r}\, d\tau.$$

These equations, respectively, express the solution of Poisson's partial differential equation (2), and that of the corresponding vector partial differential equation (4).

§52
Gravitational Potential and Field-Intensity

The potential U at a field-point $P(x, y, z)$ of a distribution of matter in a region V is defined by the equation:

(1) $$U = \int_V \frac{\delta}{r}\, dV,$$

where δ represents the density of the matter at any point $\Pi(\xi, \eta, \zeta)$ of V, and r the distance from Π to P.

We shall show that the gradient ∇U of U at the point $P(x, y, z)$ specifies the gravitational field-intensity at P attributable to the matter in V. We have:

(2) $$\nabla U = \nabla \int_V \frac{\delta}{r}\, dV = \int_V \left(\nabla \frac{1}{r}\right) \delta dV = -\int_V \frac{\mathbf{r}_0}{r^2}\, \delta dV.$$

The contribution to ∇U arising from the matter in a volume element dV at Π will, therefore, be:

$$-\frac{\mathbf{r}_0}{r^2}\, \delta dV.$$

Now, in accordance with Newton's law of gravitation, this expression also represents, to a factor of proportionality k (depending on the units),[1] the gravitational force in magnitude and direction which the matter in dV would exert upon a unit mass of matter supposed concentrated at P. Hence:

$$-k\int_V \frac{\mathbf{r}_0}{r^2}\, \delta dV$$

will represent the total gravitational force, or field-intensity, in magnitude and direction, attributable to all the matter within V. If $\mathbf{f}$ denote the field-intensity, we shall therefore have:

(3) $$\mathbf{f} = k\nabla U,$$

[1] In absolute c.g.s. units $k = 6.664 \times 10^{-8}$, numerically.

or, equivalently, in Cartesian notation:

$$(4) \qquad f_1 = k\frac{\partial U}{\partial x}, \qquad f_2 = k\frac{\partial U}{\partial y}, \qquad f_3 = k\frac{\partial U}{\partial z},$$

where f_1, f_2, f_3 are the scalar **i**, **j**, **k**-components of **f**.

The gravitational potential U must satisfy Poisson's equation at all points:

$$(5) \qquad \nabla^2 U = \text{div}\, \nabla U = -4\pi\delta.$$

Consequently, **f** must satisfy the equation:

$$(6) \qquad \text{div}\, \mathbf{f} = -4\pi k\delta.$$

Since the curl of the gradient of a scalar point function vanishes, it follows that:

$$(7) \qquad \text{curl}\, \mathbf{f} = 0,$$

throughout the gravitational field.

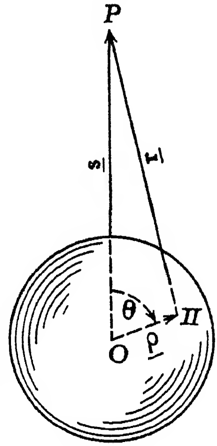


Fig. 33.

The work which would be done by the forces of the field upon a concentrated unit mass as it is made to pass from any point A to any other point B in the field furnishes a measure of the difference of gravitational potential between A and B. For, if $d\mathbf{s}$ represent an element of any path connecting A and B, then:

$$(8) \qquad \int_A^B \mathbf{f} \cdot d\mathbf{s} = \int_A^B k\nabla U \cdot d\mathbf{s} = k\int_A^B dU = k(U_B - U_A).$$

As a concrete example we shall give the solution of the problem of finding U and **f** for a gravitational field due to a material sphere of radius a. See Fig. 33.

Using polar co-ordinates (ρ, θ, ϕ) with origin at the center of the sphere, we have:

$$U = \int \frac{\delta dV}{r} = \int_0 \int_0^\pi \int_0^{2\pi} \frac{\delta \rho^2 \sin\theta d\rho d\theta d\phi}{r}.$$

In place of θ we introduce r as independent variable by means of the equation:

$$r^2 = s^2 + \rho^2 - 2s\rho\cos\theta,$$

where $s = OP$. Differentiating this equation, holding s and ρ constant, we find:

$$\frac{\rho \sin\theta d\theta}{r} = \frac{dr}{s}.$$

For the limits of r, corresponding to the limits 0 and π for θ, we have:

when $\theta = 0$, $r = s - \rho$, when $\theta = \pi$, $r = s + \rho$, for $s \geqslant a$;
when $\theta = 0$, $r = s - \rho$, when $\theta = \pi$, $r = s + \rho$, for $a \geqslant s \geqslant \rho$;
when $\theta = 0$, $r = \rho - s$, when $\theta = \pi$, $r = \rho + s$, for $a \geqslant \rho \geqslant s$.

Hence:

$$U = \int_0^a \int_{s-\rho}^{s+\rho} \int_0^{2\pi} \frac{\delta \rho d\rho dr d\phi}{s}, \qquad s \geqslant a;$$

$$U = \int_0^s \int_{s-\rho}^{s+\rho} \int_0^{2\pi} \frac{\delta \rho d\rho dr d\phi}{s} + \int_s^a \int_{\rho-s}^{\rho+s} \int_0^{2\pi} \frac{\delta \rho d\rho dr d\phi}{s}, \qquad s \leqslant a.$$

Integrating, we find:

$$U = \frac{4\pi\delta a^3}{3s} = \frac{m}{s}, \qquad s \geqslant a, \tag{9}$$

$$U = \frac{4\pi\delta}{3}s^2 + 2\pi\delta(a^2 - s^2) = -\frac{m}{2a^3}s^2 + \frac{3m}{2a}, \qquad s \leqslant a, \tag{10}$$

where m represents the total mass of the sphere. Using the values for U given by these equations in equation (3), we obtain the following expressions for **f** at points outside and at points inside the sphere:

$$\mathbf{f} = k\nabla U = -km\frac{\mathbf{s}}{s^3}, \qquad s \geqslant a; \tag{11}$$

$$\mathbf{f} = k\nabla U = -\frac{km}{a^3}\,\mathbf{s}, \qquad s \leqslant a. \tag{12}$$

From the formulas found for **f** and U it appears that at the center of the sphere:

$$\mathbf{f} = 0, \qquad U = \frac{3m}{2a}; \tag{13}$$

and that, in the region outside it, the values for **f** and U are the same as they would be if the mass of the sphere were concentrated at its center. Furthermore, inside the sphere **f** varies directly as the distance from the center, while outside it varies inversely as the square of the distance from the center.

In the treatment of this problem it has been assumed that δ is discontinuous at the surface of the sphere, thus violating one of the sufficient conditions (Art. 50) imposed upon the point-function u, in order to secure the convergence of the integral defining its potential U, and the continuity of U and its first space derivatives. But these results are all secured in the above problem (in

which δ replaces u), in spite of the assumed discontinuity in δ at the surface of the sphere, as is seen by inspection of the expressions found for U and ∇U. We have here an example of those cases for which sufficient conditions are not all necessary.

§53

The Vector Potential and the Magnetic Field-Intensity Due to a Distribution of Steady Electric Currents

Consider a distribution of steady electric currents in a system of conductors. At any source-point $\Pi(\xi, \eta, \zeta)$ in a conductor, let $\mathbf{c}$ represent the current per unit area in magnitude and direction. The vector potential $\mathbf{G}$ of the current distribution at any field-point $P(x, y, z)$ is then defined by the equation:

$$\mathbf{G} = \int \frac{\mathbf{c}}{r}\, dV, \tag{1}$$

where the integration is to be extended throughout all conductors, and where r is the distance from Π to P, and dV an element of volume.

We shall now show that the magnetic force at P, represented by $\mathbf{H}$, which is attributable to the currents, is proportional to curl $\mathbf{G}$. We have:

$$\begin{aligned} \text{curl } \mathbf{G} &= \int \text{curl}\left(\frac{\mathbf{c}}{r}\right) dV \\ &= \int \left(\nabla \frac{1}{r} \times \mathbf{c} + \frac{1}{r} \nabla \times \mathbf{c}\right) dV, \end{aligned}$$

with the aid of formula (6), Art. 39. Hence, noting that $\nabla \times \mathbf{c}$ vanishes, since $\mathbf{c}$ is independent of the co-ordinates of P:

$$\text{curl } \mathbf{G} = \int \mathbf{c} \times \frac{\mathbf{r}_0}{r^2}\, dV. \tag{2}$$

If we suppose dV to represent the volume of a very small right cylinder with axis parallel to $\mathbf{c}$ and of section ω and length l, then:

$$\mathbf{c} \times \frac{\mathbf{r}_0}{r^2}\, dV = \frac{(c\omega l) \sin (\mathbf{c}, \mathbf{r})}{r^2}\, \boldsymbol{\varepsilon}, \tag{3}$$

where $\boldsymbol{\varepsilon}$ is a unit vector perpendicular to $\mathbf{c}$ and $\mathbf{r}_0$ and in the direction of $\mathbf{c} \times \mathbf{r}_0$. The scalar value of the current in the cylindrical element of volume is $c\omega$, and this multiplied by l represents a current element $(c\omega l)$. Now, to a factor of proportionality κ (depending on the

units),[1] the contribution in magnitude and direction of the current element at Π to the total magnetic force, or field-intensity, **H** at P is, by a well known law of electromagnetism, expressed by the right-hand member of the last equation. Hence, the integral of the left-hand member of this equation throughout the conductors must be equal to $\mathbf{H}/\kappa$. It follows, then, from equation (2) that:

$$\mathbf{H} = \kappa \operatorname{curl} \mathbf{G}. \tag{4}$$

As a simple example of the use of a vector potential function we shall calculate the vector potential **G**, and from it the magnetic force **H**, due to a steady electric current, whose density is represented by **c**, flowing in a very long straight wire of infinitesimal section ω.

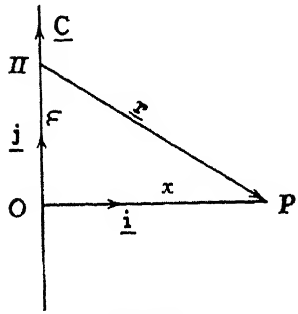


Fig. 34.

Referring to Fig. 34, let P be the field-point, and let the orthogonal projection of P upon the wire be taken as the origin O of an **i**, **j**, **k**-system of axes, **i** having the direction from O toward P, and **j** the direction of **c**. Then the co-ordinates of the field point P and of the source point Π will be x, o, o, and o, η, o, respectively. If l_2 represent the length of a portion of the wire extending from O in the positive direction of the **j**-axis, and l_1 that of a portion extending in the negative direction of this axis, then from formula (1) we shall have:

$$\mathbf{G} = \int_{-l_1}^{l_2} \frac{\mathbf{c}\omega d\eta}{r},$$

where:

$$r = \sqrt{\eta^2 + x^2};$$

and $\mathbf{c}\omega$ represents the total current flowing in the wire; denoting this by **C**, we shall then have:

$$\mathbf{C}\int_{-l_1}^{l_2} \frac{d\eta}{\sqrt{\eta^2 + x^2}}.$$

Integrating, we find:

$$\begin{aligned} \mathbf{G} &= \mathbf{C}\,[\log_e (l_2 + \sqrt{l_2^2 + x^2}) - \log_e (-l_1 + \sqrt{l_1^2 + x^2})] \\ &= \mathbf{C}\,[\log_e (l_2 + \sqrt{l_2^2 + x^2}) + \log_e (l_1 + \sqrt{l_1^2 + x^2}) - \log x^2]. \end{aligned}$$

[1] In "rational" units $\kappa = 1/4\pi c$, where c has the numerical value, 3×10^{10}, of the velocity of light in vacuum.

We now assume both l_1 and l_2 to be very large in comparison with x. Under this assumption we obtain, with close approximation:

$$G = K - 2C \log_e x, \tag{5}$$

where:

$$K = C \log_e (4l_1 l_2).$$

From equations (4) and (5) we find:

$$\mathbf{H} = \kappa \operatorname{curl} (\mathbf{K} - 2\mathbf{C} \log_e x) = -2\kappa \operatorname{curl} (\mathbf{C} \log_e x);$$

but by formula (6), Art. 39, remembering that **C** represents a constant quantity:

$$\operatorname{curl} (\mathbf{C} \log_e x) = (\nabla \log_e x) \times \mathbf{C} = \frac{1}{x}\mathbf{i} \times \mathbf{C} = \frac{C}{x}\mathbf{i} \times \mathbf{j} = \frac{C}{x}\mathbf{k}.$$

Hence:

$$\mathbf{H} = -\frac{2\kappa C}{x}\mathbf{k}. \tag{6}$$

§54
Helmholtz's Theorem

In bringing to a conclusion our brief discussion of scalar and vector potential functions, we shall give the proof of a celebrated theorem which is of importance in connection with the applications of these functions. The proof of the theorem was first given by Helmholtz in 1858 in his great paper on Vortex Motion. The theorem can be stated as follows:

Any vector point function **w** whose divergence and curl have potentials can be expressed as the sum of a lamellar part and a solenoidal part.

Let us assume it possible to express **w** in the form:

$$\mathbf{w} = \mathbf{u} + \mathbf{v}, \tag{1}$$

with:

$$\operatorname{curl} \mathbf{u} = 0, \tag{2}$$ $$\operatorname{div} \mathbf{v} = 0. \tag{3}$$

Then **u** will be a lamellar, and **v** a solenoidal vector point function, and the theorem will be established upon showing that both **u** and **v** can be found when **w** is everywhere given.

Since we assume equations (2) and (3) to be satisfied, we can take:

$$\mathbf{u} = -\nabla U, \tag{4}$$ $$\mathbf{v} = \operatorname{curl} \mathbf{G}, \tag{5}$$

where U is a scalar point function, and $\mathbf{G}$ is a solenoidal vector point function, so that:

$$\text{(6)} \qquad \operatorname{div} \mathbf{G} = 0.$$

From equations (1), (3), and (4) we find:

$$\nabla^2 U = -\operatorname{div} \mathbf{w}.$$

This is Poisson's equation (2), Art. 51, with div $\mathbf{w}$ in place of $4\pi u$; and U is therefore a scalar potential function which, by equation (1), Art. 50, is given by the equation:

$$\text{(7)} \qquad U = \frac{1}{4\pi}\int \frac{\operatorname{div} \mathbf{w}}{r}\, d\tau.$$

Hence U, and therefore $\mathbf{u} = -\nabla U$, can be found when $\mathbf{w}$ is everywhere given.

From equations (1), (2), and (5) we find:

$$\operatorname{curl}^2 \mathbf{G} = \operatorname{curl} \mathbf{w}.$$

But, by formula (10), Art. 39, and by equation (6):

$$\operatorname{curl}^2 \mathbf{G} = -\nabla^2 \mathbf{G}.$$

Hence:

$$\nabla^2 \mathbf{G} = -\operatorname{curl} \mathbf{w}.$$

This is Poisson's equation (4), Art. 51, with curl $\mathbf{w}$ in place of $4\pi\mathbf{g}$; and $\mathbf{G}$ is, therefore, a vector potential function which, by equation (2), Art. 50, is given by the equation:

$$\text{(8)} \qquad \mathbf{G} = \frac{1}{4\pi}\int \operatorname{curl} \mathbf{w}\; d\tau.$$

Hence $\mathbf{G}$, and therefore $\mathbf{v}(=\operatorname{curl} \mathbf{G})$, can be found when $\mathbf{w}$ is everywhere given.

Since it has been shown that both $\mathbf{u}$ and $\mathbf{v}$ can be found when $\mathbf{w}$ is everywhere given, the theorem is established.

As a corollary to the theorem we have the statement:

The vector point function $\mathbf{w}$ is completely determined when its divergence and curl are everywhere given.

EXERCISES ON CHAPTER V

1. From the results found in Art. 52 show directly that the gravitational potential due to a sphere of matter of uniform density satisfies Poisson's equation at points inside the sphere and Laplace's equation at points outside.

2. Find the gravitational potential of a very thin straight wire of linear density γ; then show that the magnitude f of the gravitational force at a generic point P is given by the equation:

$$f = \frac{2k\gamma}{p} \sin \frac{1}{2}\theta,$$

where p is the distance of P from the wire, θ the angle between lines drawn from P to its extremities, and k a factor of proportionality depending upon the units; show also that the direction of the force is along a line bisecting the angle θ; finally, show that the equipotential surfaces and the lines of force are, respectively, ellipsoids and hyberbolos with the ends of the wire as foci.

3. If a and b are the inner and outer radii of a spherical shell of matter of uniform density δ, $\mathbf{s}$ a line-vector to a generic point P from the center of the shell, and U the gravitational potential due to the shell, show that the data in the following table are valid:

	$s < a$	$a < s < b$	$b < s$
U	$2\pi\delta(b^2 - a^2)$,	$2\pi\delta\left(b^2 - \frac{s^2}{3}\right) - \frac{4\pi\delta}{3s} a^3$,	$\frac{4\pi\delta}{3s}(b^3 - a^3)$,
∇U	0,	$\frac{4\pi\delta}{3}\left(\frac{a^3}{s^3} - 1\right)\mathbf{s}$,	$-\frac{4\pi\delta}{3s^3}(b^3 - a^3)\,\mathbf{s}$.

4. If S is an equipotential surface for a distribution of matter, show that the gravitational potential due to it at any point outside S is the same as that due to a distribution of matter over S of surface density:

$$\mu = \frac{1}{4\pi}\mathbf{n}\cdot\mathbf{f},$$

where $\mathbf{n}\cdot\mathbf{f}$ represents the outward normal component of the gravitational force due to the original distribution; show also that the total mass of the imaginary surface distribution is equal to the actual mass within S.

5. If the gravitational potential functions U and U' due, respectively, to two systems of matter outside a closed surface have the same values at all points on the surface, show that they will be equal at all points inside the surface.

6. A distribution of electricity spherically symmetrical with respect to a point O possesses a potential which at a distance s from O is equal to ke^{-s^2}, where k is a constant; show that the law of variation of the density ρ with s is given by the equation:

$$\rho = -ke^{-s^2}(2s^2 - 3)/2\pi.$$

7. Write down the integral expression for the vector potential at a generic point P due to a steady current $\mathbf{c}$ flowing in a very thin circular

turn of wire of radius a. By taking the curl of this expression find an integral expressing the magnetic force at P due to the current, and from it show that the magnitude H of the magnetic force at a point on the axis of the turn at a distance s from its center is given by the equation:

$$H = \frac{2\pi k a^2 c}{(s^2 + a^2)^{\frac{3}{2}}},$$

where k is a constant depending upon the units.

CHAPTER VI

LINEAR VECTOR FUNCTIONS AND DYADICS

§55

Definition of a Linear Vector Function

We consider a vector $\mathbf{f}$ which stands in relationship to a second vector $\mathbf{v}$ such that, when a value for $\mathbf{v}$ is specified, a corresponding value of $\mathbf{f}$ will be determined. Such a relationship may be expressed by writing:

$$\mathbf{f} = f(\mathbf{v}),$$

and $\mathbf{f}$ is said to be a vector function of $\mathbf{v}$.

Let us now suppose $\mathbf{v}$ to be expressed in terms of its components on a base-system, determined by any three non-coplanar vectors $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$, as follows:

$$\mathbf{v} = v_1\mathfrak{a}_1 + v_2\mathfrak{a}_2 + v_3\mathfrak{a}_3,$$

where v_1, v_2, v_3 are the measure-numbers of the components. Furthermore, let us suppose that $\mathbf{f}$ is continuous and that the relationship of $\mathbf{f}$ to $\mathbf{v}$ is such that:

$$\begin{aligned} \mathbf{f} = f(\mathbf{v}) &= f(v_1\mathfrak{a}_1 + v_2\mathfrak{a}_2 + v_3\mathfrak{a}_3) \\ &= v_1 f(\mathfrak{a}_1) + v_2 f(\mathfrak{a}_2) + v_3 f(\mathfrak{a}_3); \end{aligned} \tag{1}$$

in this case $\mathbf{f}$ is said to be a Linear Vector Function of $\mathbf{v}$.

The quantities

$$f(\mathfrak{a}_1), \quad f(\mathfrak{a}_2), \quad f(\mathfrak{a}_3)$$

are evidently vectors representing the values which $\mathbf{f}$ assumes when $\mathbf{v}$ is given the values $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$. When these quantities are known, the linear vector function $f(\mathbf{v})$ will be completely determined. For, when any value of $\mathbf{v}$ is given, through specification of particular values of v_1, v_2, v_3, the corresponding value of $\mathbf{f}$ will be given by equation (1). Hence:

A linear vector function $f(\mathbf{v})$ is completely determined when its values for any three non-coplanar values of $\mathbf{v}$ are specified.

It is evident that the specification of three values for $f(\mathbf{v})$ will require in general nine parameters; for the specification requires a

knowledge of the magnitudes aud directions of three vectors which in general will be non-coplanar. The numerical values of these nine parameters will of course depend upon the base-system used in the specification.

The simplest case of a linear vector function is one for which but one parameter is required for its specification. For example, if

$$\mathbf{f} = f(\mathbf{v}) = k\mathbf{v},$$

where k is a constant, then $\mathbf{f}$ is a linear vector function of $\mathbf{v}$ in which the single parameter k is a factor of proportionality. In this case for any given value of $\mathbf{v}$ the corresponding value of $\mathbf{f}$ will represent a vector k times as great as $\mathbf{v}$.

A case in which three parameters are required for the specification of a linear vector function is presented by the equation:

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r},$$

in which $\mathbf{v}$ specifies the velocity of a point in a uniformly rotating rigid body, $\boldsymbol{\omega}$ the uniform angular velocity of the body, and $\mathbf{r}$ the position-vector of the point. Here, $\mathbf{v}$ is a linear vector function of $\mathbf{r}$ which requires three parameters for its specification, since three parameters are necessary for the specification of $\boldsymbol{\omega}$.

Physics furnishes numerous examples of linear vector functions, some of which will be considered in detail further on.

§56

A Linear Vector Function Represents an Affine Transformation

Let the vectors $\mathbf{r}$ and $\mathbf{q}$ represent the position vectors of two points P and Q, and let them be expressed in terms of their components on a base-system, determined by any three non-coplanar vectors, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, as follows:

(1) $$\mathbf{r} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3,$$

(2) $$\mathbf{q} = y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + y_3\mathbf{a}_3,$$

where the measure numbers x_1, x_2, x_3 and y_1, y_2, y_3 of $\mathbf{r}$ and $\mathbf{q}$ may be regarded as the co-ordinates of the points P and Q.

Supposing $\mathbf{q}$ to be a linear vector function of $\mathbf{r}$, we can write:

$$\mathbf{q} = x_1 f(\mathbf{a}_1) + x_2 f(\mathbf{a}_2) + x_3 f(\mathbf{a}_3);$$

or:

(3) $$\mathbf{q} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3,$$

where:

$$(4) \qquad \mathbf{b}_1 = f(\mathfrak{a}_1), \qquad \mathbf{b}_2 = f(\mathfrak{a}_2), \qquad \mathbf{b}_3 = f(\mathfrak{a}_3).$$

The vectors determined by equations (4) can be expressed in terms of their components on the base-system by the equations:

$$(5) \qquad \begin{aligned} \mathbf{b}_1 &= a_{11}\mathfrak{a}_1 + a_{21}\mathfrak{a}_2 + a_{31}\mathfrak{a}_3, \\ \mathbf{b}_2 &= a_{12}\mathfrak{a}_1 + a_{22}\mathfrak{a}_2 + a_{32}\mathfrak{a}_3, \\ \mathbf{b}_3 &= a_{13}\mathfrak{a}_1 + a_{23}\mathfrak{a}_2 + a_{33}\mathfrak{a}_3, \end{aligned}$$

where the a-coefficients are measure-numbers of the components. From equations (3) and (5) we shall then have:

$$(6) \qquad \begin{aligned} \mathbf{q} = \; & (a_{11}x_1 + a_{12}x_2 + a_{13}x_3)\,\mathfrak{a}_1 \\ & + (a_{21}x_1 + a_{22}x_2 + a_{23}x_3)\,\mathfrak{a}_2 \\ & + (a_{31}x_1 + a_{32}x_2 + a_{33}x_3)\,\mathfrak{a}_3. \end{aligned}$$

By comparison of equations (2) and (6) we find:

$$(7) \qquad \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ y_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{aligned}$$

These homogeneous linear equations express an Affine Transformation with origin fixed of the point $P(x_1, x_2, x_3)$ into the point $Q(y_1, y_2, y_3)$. They show that in such transformations finite points remain finite, and infinite points remain infinite; also that straight lines remain straight, and that parallel lines remain parallel.

The transformation is characterized by the nine constant scalar a-coefficients which, when known, determine the linear vector function $\mathbf{q}$. We may therefore say that a linear vector function represents an affine transformation with origin fixed.

For the determinant Δ of the transformation we have:

$$(8) \qquad \Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

If this determinant does not vanish, the transformation and the corresponding linear vector function $\mathbf{q}$ which represents it are said to be non-degenerate. In this case the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ will be non-coplanar, since the condition

$$\mathbf{b}_1 \cdot \mathbf{b}_2 \times \mathbf{b}_3 \neq 0$$

will be satisfied; for, upon forming the scalar triple product of these vectors, as given by equations (5), we find:

$$\mathbf{b}_1 \cdot \mathbf{b}_2 \times \mathbf{b}_3 = \Delta \mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3,$$

and the scalar triple product on the right cannot vanish, since $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ are non-coplanar by hypothesis.

Equations (7) might be used as equations of definition for a linear vector function.[1] Starting with them as equations specifying the measure-numbers of the components of a vector $\mathbf{q}$ in terms of those of a second vector $\mathbf{v}$ on a given base-system, it can easily be shown that $\mathbf{q}$ must be a linear vector function of $\mathbf{v}$ as defined in the preceding article.

With the aid of equations (7) the following properties of a linear vector function can easily be found:

$$f(k\mathbf{v}) = kf(\mathbf{v}), \tag{9}$$

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}), \tag{10}$$

where k represents any pure number, and where $\mathbf{u}$ and $\mathbf{v}$ represent any two vectors.

The property expressed by the last equation can also be used as a defining property of a linear vector function.[1]

§57

Definition of a Dyadic

The Direct Products of a Dyadic and a Vector

The theory of linear vector functions is most conveniently developed with the aid of certain operators known as Dyadics. The theory of dyadic operators is essentially equivalent to that of linear vector functions, as will be seen.

Two vectors placed in juxtaposition with neither $\cdot$ nor $\times$ between constitute what is called a Dyad. For example:

$$\mathbf{AB}$$

is a dyad.

Any polynomial of dyads, such as

$$\mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{A}_n\mathbf{B}_n,$$

is called a Dyadic.

The first vector of a dyad is called its Antecedent and the second its Consequent. The Antecedents and the Consequents of a dyadic are, respectively, the antecedents and the consequents of its constituent dyads.

In general a dyadic will be denoted by a bold large Greek letter, the letter $\boldsymbol{\Phi}$ being most frequently used for this purpose.

[1] Cf. Gibbs-Wilson, Vector Analysis, p. 262.

A dyad is said to be multiplied by a scalar k when either of its factors is multiplied by k. A dyadic $\mathbf{\Phi}$ is said to be multiplied by a scalar k when each of its dyads is multiplied by k; and the product is denoted by $k\mathbf{\Phi}$ or $\mathbf{\Phi}k$.

If the order of the vectors in each dyad of a dyadic $\mathbf{\Phi}$ be reversed, a dyadic called the Conjugate of $\mathbf{\Phi}$ is thereby obtained. The conjugate of the conjugate of a dyadic is, of course, the dyadic itself.

The Direct Products of a dyadic

$$\mathbf{\Phi} = \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{A}_n\mathbf{B}_n$$

and a vector $\mathbf{v}$ are denoted by $\mathbf{\Phi} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{\Phi}$, and are defined as follows:

$$\mathbf{\Phi} \cdot \mathbf{v} = \mathbf{A}_1\mathbf{B}_1 \cdot \mathbf{v} + \mathbf{A}_2\mathbf{B}_2 \cdot \mathbf{v} + \cdots + \mathbf{A}_n\mathbf{B}_n \cdot \mathbf{v},$$
$$\mathbf{v} \cdot \mathbf{\Phi} = \mathbf{v} \cdot \mathbf{A}_1\mathbf{B}_1 + \mathbf{v} \cdot \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{v} \cdot \mathbf{A}_n\mathbf{B}_n.$$

In the first of these equations the dyadic $\mathbf{\Phi}$ is said to act as a Prefactor, and in the second as a Postfactor. It will be noted that a dyadic acting as prefactor or postfactor upon a vector produces a vector; furthermore, that a dyadic acting as prefactor upon a vector produces the same vector as its conjugate acting as postfactor, and vice versa.

It follows from the last two equations that:

$$\mathbf{\Phi} \cdot (k\mathbf{v}) = k\mathbf{\Phi} \cdot \mathbf{v},$$
$$(k\mathbf{v}) \cdot \mathbf{\Phi} = k\mathbf{v} \cdot \mathbf{\Phi},$$

where k is any scalar.

We shall now show that:

A dyadic acting as prefactor or postfactor upon any vector produces a linear vector function of this vector.

In the expression for $\mathbf{\Phi} \cdot \mathbf{v}$ just given, upon replacing the vector $\mathbf{v}$ by the sum of its three components on an arbitrary base-system determined by three non-coplanar vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, so that

$$\mathbf{v} = v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3,$$

where v_1, v_2, v_3, are the measure-numbers of the components, we find:

$$\begin{aligned}\mathbf{\Phi} \cdot \mathbf{v} &= \mathbf{\Phi} \cdot (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3)\\ &= v_1\mathbf{\Phi} \cdot \mathbf{a}_1 + v_2\mathbf{\Phi} \cdot \mathbf{a}_2 + v_3\mathbf{\Phi} \cdot \mathbf{a}_3.\end{aligned}$$

Upon comparison of this equation with the defining equation (1), Art. 55, for a linear vector function, viz:

$$\begin{aligned}f(\mathbf{v}) &= f(v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3)\\ &= v_1 f(\mathbf{a}_1) + v_2 f(\mathbf{a}_2) + v_3 f(\mathbf{a}_3),\end{aligned}$$

it is seen that $\mathbf{\Phi} \cdot \mathbf{v}$ represents a linear vector function of $\mathbf{v}$ for which:

$$f(\mathbf{a}_1) = \mathbf{\Phi} \cdot \mathbf{a}_1, \qquad f(\mathbf{a}_2) = \mathbf{\Phi} \cdot \mathbf{a}_2, \qquad f(\mathbf{a}_3) = \mathbf{\Phi} \cdot \mathbf{a}_3.$$

In like manner $\mathbf{v} \cdot \mathbf{\Phi}$ can also be shown to represent a linear vector function of $\mathbf{v}$ for which:

$$f(\mathbf{a}_1) = \mathbf{a}_1 \cdot \mathbf{\Phi}, \qquad f(\mathbf{a}_2) = \mathbf{a}_2 \cdot \mathbf{\Phi}, \qquad f(\mathbf{a}_3) = \mathbf{a}_3 \cdot \mathbf{\Phi}.$$

§58

Criteria of Equality for Two Dyadics

Two dyadics $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are said to be equal if they satisfy the condition:

(1) $$\mathbf{u} \cdot \mathbf{\Phi} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{\Psi} \cdot \mathbf{v}, \text{ for all values of } \mathbf{u} \text{ and } \mathbf{v}.$$

The products in both members of this expression are associative, as is easily seen by substituting for $\mathbf{\Phi}$ or $\mathbf{\Psi}$ its general expression as a sum of dyads, and multiplying out. The condition (1) is equivalent to either one of the following conditions:

(2) $$\mathbf{\Phi} \cdot \mathbf{v} = \mathbf{\Psi} \cdot \mathbf{v}, \text{ for all values of } \mathbf{v},$$

(3) $$\mathbf{u} \cdot \mathbf{\Phi} = \mathbf{u} \cdot \mathbf{\Psi}, \text{ for all values of } \mathbf{u}.$$

For, since the products in condition (1) are associative, we can write:

$$\mathbf{u} \cdot (\mathbf{\Phi} \cdot \mathbf{v}) = \mathbf{u} \cdot (\mathbf{\Psi} \cdot \mathbf{v}), \text{ for all values of } \mathbf{u} \text{ and } \mathbf{v};$$

but this requires that condition (2) shall also be satisfied, and conversely, if condition (2) is satisfied, condition (1) must also be satisfied; hence (1) and (2) are equivalent; in like manner conditions (1) and (3) can be shown to be equivalent. Hence, all three conditions are equivalent.

In fact, two dyadics $\mathbf{\Phi}$ and $\mathbf{\Psi}$ will be equal, provided equations (1), (2), and (3) are satisfied for any three non-coplanar values of $\mathbf{u}$ and $\mathbf{v}$. For, as seen above, a dyadic $\mathbf{\Phi}$ or $\mathbf{\Psi}$ acting as prefactor or postfactor upon a vector $\mathbf{v}$ determines a linear vector function of $\mathbf{v}$, and, as seen in Art. 55, this function is completely determined when its values for any three non-coplanar values of $\mathbf{v}$ are known.

We can therefore state that:

A dyadic $\mathbf{\Phi}$ will be completely determined when the values of $\mathbf{\Phi} \cdot \mathbf{v}$, or of $\mathbf{v} \cdot \mathbf{\Phi}$, for any three non-coplanar values of $\mathbf{v}$ are known.

The following statement is easily proved:

Any linear vector function $f(\mathbf{v})$ can be produced by a dyadic $\mathbf{\Phi}$ acting upon $\mathbf{v}$ as a prefactor, or by its conjugate $\mathbf{\Phi}_c$, acting upon $\mathbf{v}$ as a postfactor:

It was shown in Art. 55 that a linear vector function $f(\mathbf{v})$ is completely determined when its values for any three non-coplanar values of $\mathbf{v}$ viz: $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, are known. Let $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ denote the values of the function when $\mathbf{v}$ has the values $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, so that:

$$f(\mathbf{a}_1) = \mathbf{f}_1, \qquad f(\mathbf{a}_2) = \mathbf{f}_2, \qquad f(\mathbf{a}_3) = \mathbf{f}_3.$$

Then the linear vector function will evidently be represented by a dyadic $\mathbf{\Phi}$ acting as a prefactor if:

$$\mathbf{\Phi} = \mathbf{f}_1\mathbf{a}^1 + \mathbf{f}_2\mathbf{a}^2 + \mathbf{f}_3\mathbf{a}^3,$$

where $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ are the vectors of the system reciprocal to the $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$-system. It will also be represented by the conjugate of $\mathbf{\Phi}$, viz:

$$\mathbf{\Phi}_c = \mathbf{a}^1\mathbf{f}_1 + \mathbf{a}^2\mathbf{f}_2 + \mathbf{a}^3\mathbf{f}_3,$$

acting as a postfactor.

From the criteria of equality of two dyadics it follows that:

The distributive law of multiplication is valid in the expansion of compound dyads, provided the order of the vectors be maintained.

It will suffice to verify this statement for the special case:

$$(\mathbf{a} + \mathbf{b})(\mathbf{c} + \mathbf{d}) = \mathbf{ac} + \mathbf{ad} + \mathbf{bc} + \mathbf{bd}. \tag{4}$$

We have:

$$\begin{aligned}(\mathbf{a} + \mathbf{b})(\mathbf{c} + \mathbf{d}) \cdot \mathbf{v} &= (\mathbf{a} + \mathbf{b})(\mathbf{c} \cdot \mathbf{v} + \mathbf{d} \cdot \mathbf{v}) \\ &= \mathbf{ac} \cdot \mathbf{v} + \mathbf{ad} \cdot \mathbf{v} \\ &\quad + \mathbf{bc} \cdot \mathbf{v} + \mathbf{bd} \cdot \mathbf{v} \\ &= (\mathbf{ac} + \mathbf{ad} + \mathbf{bc} + \mathbf{bd}) \cdot \mathbf{v},\end{aligned}$$

for all values of $\mathbf{v}$. Hence, equation (4) is verified.

§59

Reduction of a Dyadic to a Tri-nomial Form

It has been shown above that a dyadic acting as a prefactor upon a vector $\mathbf{v}$ produces a linear vector function of $\mathbf{v}$ and, conversely, that any linear vector function of a vector $\mathbf{v}$ can be produced by an appropriate dyadic acting as a prefactor upon $\mathbf{v}$. Now, since, in the general case, a linear vector function requires nine parameters for its specification, it follows that, in the general case, a dyadic will also require but nine parameters for its specification. The dyadic, defined in Art. 57 as a polynomial consisting of an arbitrary number of dyads, must, therefore, be reducible to a form for which nine parameters suffice for its specification.

Consider, then, the dyadic $\mathbf{\Phi}$ expressed in the polynomial form:

$$\mathbf{\Phi} \quad \sum_{i=1}^{n} \mathbf{A}_i\mathbf{B}_i.$$

The antecedent of each dyad in the sum on the right can be expressed in terms of its components on a base-system determined by any three non-coplanar vectors, $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ as follows:

$$\mathbf{A}_i = \sum_{j=1}^{3} A_{ij}\mathfrak{a}_j.$$

Remembering the distributive law of multiplication stated in the preceding article, we can now write:

$$\mathbf{\Phi} = \sum_{i=1} \mathbf{A}_i\mathbf{B}_i = \sum_{i=1}^{n} \sum_{j=1}^{3} A_{ij}\mathfrak{a}_j\mathbf{B} = \sum_{j=1}^{3} \mathfrak{a}_j \sum_{i=1}^{n} A_{ij}\mathbf{B}_i.$$

And, if we let:

$$\mathbf{b}_j = \sum_{i=1}^{n} A_{ij}\mathbf{B}_i, \qquad (j = 1, 2, 3),$$

then:

$$\mathbf{\Phi} = \sum_{i=1}^{n} \mathbf{A}_i\mathbf{B}_i = \sum_{j=1}^{3} \mathfrak{a}_j\mathbf{b}_j = \mathfrak{a}_1\mathbf{b}_1 + \mathfrak{a}_2\mathbf{b}_2 + \mathfrak{a}_3\mathbf{b}_3.$$

In a similar manner the reduction might be made to a trinomial form in which the consequents are chosen arbitrarily instead of the antecedents.

Any dyadic $\mathbf{\Phi}$ can, therefore, be reduced to a form consisting of but three dyads whose antecedents or consequents may be chosen as any three non-coplanar vectors. If, in the reduced form, the antecedents are non-coplanar and also the consequents, the dyadic is called a Complete Dyadic.

Since the three antecedents (or consequents) in the trinomial form for $\mathbf{\Phi}$ can be taken as any three non-coplanar vectors, the three consequents (or antecedents) must determine $\mathbf{\Phi}$ completely. Hence, in general, nine parameters are required for the specification of a complete dyadic, as was foreseen.

When and only when the dyadic $\mathbf{\Phi}$ is such that the antecedents (or consequents) in the reduced trinomial form are coplanar, it

can be reduced to a form consisting of two terms only. If incapable of further reduction, it is then called a Planar Dyadic.

When and only when the dyadic Φ is such that after reduction to the trinomial form, the antecedents (or consequents) are collinear, it can be further reduced to a dyadic consisting of but one term. It is then called a Linear Dyadic.

The proofs of the theorems in the last two paragraphs may be supplied by the reader as an exercise, or they may be found in the Vector Analysis of Gibbs-Wilson, pp. 282–283.

Planar and linear dyadics possess the following obvious properties:

A planar dyadic, acting as a prefactor upon any vector $\mathbf{v}$, will produce a linear vector function of $\mathbf{v}$ coplanar with its antecedents; acting as a postfactor, it will produce a linear vector function of $\mathbf{v}$ coplanar with its consequents.

A linear dyadic, acting as a prefactor upon any vector $\mathbf{v}$, will produce a linear vector function of $\mathbf{v}$ collinear with its antecedent; acting as a postfactor, it will produce a linear vector function of $\mathbf{v}$ collinear with its consequent.

By the method given above for the reduction of a dyadic to a trinomial form, any two dyadics can be reduced to trinomial forms for which the antecedents (or consequents) are respectively the same; by the criteria of equality for two dyadics (Art. 58), if the consequents (or antecedents) are also respectively the same, the two dyadics will be equal; conversely, two equal dyadics can be reduced to trinomial forms in which their antecedents and consequents are respectively the same.

§60

The Nonian Form of a Dyadic

Let us suppose any dyadic Φ to have been reduced to the trinomial form:

$$\Phi = \mathfrak{a}_1\mathbf{b}_1 + \mathfrak{a}_2\mathbf{b}_2 + \mathfrak{a}_3\mathbf{b}_3,$$

in which the antecedents $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ are arbitrary non-coplanar vectors. By expressing the consequents $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ in terms of their components on a base-system determined by three arbitrary non-coplanar vectors, $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$, and remembering the distributive law for dyad products, Φ can be expressed in the following form, involving nine dyads, the antecedents of which are arbitrary non-coplanar vectors, and also the consequents:

$$
\begin{aligned}
(1) \qquad \boldsymbol{\Phi} = \ & a_{11}\mathbf{a}_1\mathbf{c}_1 + a_{12}\mathbf{a}_1\mathbf{c}_2 + a_{13}\mathbf{a}_1\mathbf{c}_3 \\
& + a_{21}\mathbf{a}_2\mathbf{c}_1 + a_{22}\mathbf{a}_2\mathbf{c}_2 + a_{23}\mathbf{a}_2\mathbf{c}_3 \\
& + a_{31}\mathbf{a}_3\mathbf{c}_1 + a_{32}\mathbf{a}_3\mathbf{c}_2 + a_{33}\mathbf{a}_3\mathbf{c}_3,
\end{aligned}
$$

where the a's are parametric coefficients.

If, in particular, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and also $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ are taken, respectively, as the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, then:

$$
\begin{aligned}
(2) \qquad \boldsymbol{\Phi} = \ & a_{11}\mathbf{ii} + a_{12}\mathbf{ij} + a_{13}\mathbf{ik} \\
& + a_{21}\mathbf{ji} + a_{22}\mathbf{jj} + a_{23}\mathbf{jk} \\
& + a_{31}\mathbf{ki} + a_{32}\mathbf{kj} + a_{33}\mathbf{kk}.
\end{aligned}
$$

This is called the Nonian Form of the dyadic $\boldsymbol{\Phi}$. In it are represented all of the nine dyads formable with the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.

Two dyadics will be equal if, when expressed in nonian form, the coefficients of corresponding dyads are equal.

Conversely, if the coefficients of corresponding dyads are equal in the nonian forms of two dyadics, then they will be equal.

The truth of these statements follows directly from the criteria of equality for two dyadics.

§61

Symmetric and Anti-Symmetric Dyadics

A dyadic which is identical with its conjugate is called a Symmetric Dyadic. In a symmetric dyadic the antecedent and consequent of each of its dyads can be interchanged.

A dyadic $\boldsymbol{\Phi}$ and its conjugate $\boldsymbol{\Phi}_c$ can be expressed in the nonian forms:

$$
\begin{aligned}
(1) \qquad \boldsymbol{\Phi} = \ & a_{11}\mathbf{ii} + a_{12}\mathbf{ij} + a_{13}\mathbf{ik} \\
& + a_{21}\mathbf{ji} + a_{22}\mathbf{jj} + a_{23}\mathbf{jk} \\
& + a_{31}\mathbf{ki} + a_{32}\mathbf{kj} + a_{33}\mathbf{kk},
\end{aligned}
\qquad
\begin{aligned}
(2) \qquad \boldsymbol{\Phi}_c = \ & a_{11}\mathbf{ii} + a_{12}\mathbf{ji} + a_{13}\mathbf{ki} \\
& + a_{21}\mathbf{ij} + a_{22}\mathbf{jj} + a_{23}\mathbf{kj} \\
& + a_{31}\mathbf{ik} + a_{32}\mathbf{jk} + a_{33}\mathbf{kk}.
\end{aligned}
$$

With the aid of the criteria of equality for two dyadics, it is easily seen that $\boldsymbol{\Phi} = \boldsymbol{\Phi}_c$ if and only if:

$$(3) \qquad a_{12} = a_{21}, \qquad a_{23} = a_{32}, \qquad a_{31} = a_{13}.$$

If the dyadic $\boldsymbol{\Phi}$ is symmetric, it can, therefore, be expressed as follows:

$$
\begin{aligned}
(4) \qquad \boldsymbol{\Phi} = \ & a_{11}\mathbf{ii} + a_{22}\mathbf{jj} + a_{33}\mathbf{kk} \\
& + a_{12}(\mathbf{ij} + \mathbf{ji}) + a_{23}(\mathbf{jk} + \mathbf{kj}) + a_{31}(\mathbf{ki} + \mathbf{ik}).
\end{aligned}
$$

It will be shown later (Art. 68) that any symmetric dyadic $\mathbf{\Phi}$ can be reduced to the form:

$$\mathbf{\Phi} = a_{11}\mathbf{ii} + a_{22}\mathbf{jj} + a_{33}\mathbf{kk}, \tag{5}$$

by a special choice of the **i, j, k**-system of unit vectors.

A symmetric dyadic is sometimes called a Self-Conjugate Dyadic.

A dyadic is said to be Anti-Symmetric when it is identical with the negative of its conjugate.

From the expressions for $\mathbf{\Phi}$ and $\mathbf{\Phi}_c$ in nonian forms, given by equations (1) and (2), it can easily be seen with the aid of the criterion of equality of two dyadics that $\mathbf{\Phi} = -\mathbf{\Phi}_c$ if and only if:

$$a_{11} = a_{22} = a_{33} = 0, \quad a_{12} = -a_{21}, \quad a_{23} = -a_{32}, \quad a_{31} = -a_{13}. \tag{6}$$

An anti-symmetric dyadic $\mathbf{\Phi}$ can, therefore, be put in the form:

$$\mathbf{\Phi} = a_{12}(\mathbf{ij} - \mathbf{ji}) + a_{23}(\mathbf{jk} - \mathbf{kj}) + a_{31}(\mathbf{ki} - \mathbf{ik}). \tag{7}$$

Any anti-symmetric dyadic must be a planar dyadic, as can be shown as follows:

Consider a dyadic $\mathbf{\Phi}$ and its conjugate $\mathbf{\Phi}_c$ expressed in the trinomial forms:

$$\mathbf{\Phi} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3,$$
$$\mathbf{\Phi}_c = \mathbf{b}_1\mathbf{a}_1 + \mathbf{b}_2\mathbf{a}_2 + \mathbf{b}_3\mathbf{a}_3.$$

If $\mathbf{\Phi}$ is anti-symmetric, then:

$$\begin{aligned}\mathbf{\Phi}\cdot\mathbf{v} = \frac{1}{2}(\mathbf{\Phi} - \mathbf{\Phi}_c)\cdot\mathbf{v} = \frac{1}{2}[(&\mathbf{a}_1\mathbf{b}_1\cdot\mathbf{v} - \mathbf{b}_1\mathbf{a}_1\cdot\mathbf{v})\\ &+ (\mathbf{a}_2\mathbf{b}_2\cdot\mathbf{v} - \mathbf{b}_2\mathbf{a}_2\cdot\mathbf{v})\\ &+ (\mathbf{a}_3\mathbf{b}_3\cdot\mathbf{v} - \mathbf{b}_3\mathbf{a}_3\cdot\mathbf{v})],\end{aligned}$$

where $\mathbf{v}$ is an arbitrary vector. Now:

$$\begin{aligned}\mathbf{a}_1\mathbf{b}_1\cdot\mathbf{v} - \mathbf{b}_1\mathbf{a}_1\cdot\mathbf{v} &= -(\mathbf{a}_1\times\mathbf{b}_1)\times\mathbf{v},\\ \mathbf{a}_2\mathbf{b}_2\cdot\mathbf{v} - \mathbf{b}_2\mathbf{a}_2\cdot\mathbf{v} &= -(\mathbf{a}_2\times\mathbf{b}_2)\times\mathbf{v},\\ \mathbf{a}_3\mathbf{b}_3\cdot\mathbf{v} - \mathbf{b}_3\mathbf{a}_3\cdot\mathbf{v} &= -(\mathbf{a}_3\times\mathbf{b}_3)\times\mathbf{v};\end{aligned}$$

and, if we introduce a new vector $\mathbf{\Phi}_\times$, called the Vector of $\mathbf{\Phi}$, formed by inserting a $\times$ between the two vectors of each dyad of $\mathbf{\Phi}$, then:

$$\mathbf{\Phi}_\times = \mathbf{a}_1\times\mathbf{b}_1 + \mathbf{a}_2\times\mathbf{b}_2 + \mathbf{a}_3\times\mathbf{b}_3, \tag{8}$$

and $\Phi \cdot \mathbf{v}$ can be expressed in the form:

$$(9) \qquad \Phi \cdot \mathbf{v} = -\frac{1}{2}\Phi_\times \times \mathbf{v}.$$

Therefore, the anti-symmetric dyadic Φ, acting as a prefactor upon any vector $\mathbf{v}$, produces a linear vector function $\Phi \cdot \mathbf{v}$ which is coplanar with respect to a plane perpendicular to the vector of Φ; furthermore, $\Phi \cdot \mathbf{v}$ is perpendicular to the vector $\mathbf{v}$ itself. It is, therefore, permissible to write:

$$(10) \qquad \Phi = \mathbf{a}\mathbf{b} - \mathbf{b}\mathbf{a},$$

where $\mathbf{a}$ and $\mathbf{b}$ are two non-collinear vectors in a plane perpendicular to the vector $\Phi_\times$. In particular, if this plane be taken as the $\mathbf{i}$, $\mathbf{j}$-plane, the vectors $\mathbf{a}$ and $\mathbf{b}$ may be taken collinear with $\mathbf{i}$ and $\mathbf{j}$ respectively, and in this case we may write:

$$(11) \qquad \Phi = a_{12}(\mathbf{i}\mathbf{j} - \mathbf{j}\mathbf{i}),$$

where a_{12} is a constant. Each of the last two forms for Φ represents it as a planar dyadic.

It follows that, by a suitable choice of the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system of axes, the general form (7) for an anti-symmetric dyadic can be reduced to the special form (11); or, in other words, if the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system of axes be suitably chosen, it is possible to make the coefficients a_{23} and a_{31} of the general form (7) equal to zero.

An anti-symmetric dyadic is sometimes called an Anti-Self-Conjugate Dyadic.

§62
The Idemfactor or Unit Dyadic

A dyadic which, acting as prefactor or postfactor upon a vector, produces the vector itself, is called an Idemfactor or Unit Dyadic.

General forms for the idemfactor are:

$$(1) \quad \mathbf{I} = \mathbf{a}^1\mathbf{a}_1 + \mathbf{a}^2\mathbf{a}_2 + \mathbf{a}^3\mathbf{a}_3; \qquad (2) \quad \mathbf{I} = \mathbf{a}_1\mathbf{a}^1 + \mathbf{a}_2\mathbf{a}^2 + \mathbf{a}_3\mathbf{a}^3,$$

where $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are arbitrary non-coplanar vectors, and $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ are the vectors of their reciprocal system. For, as shown in Art. 20, any vector $\mathbf{v}$ can be expressed in the forms:

$$\mathbf{v} = \mathbf{a}^1\mathbf{a}_1 \cdot \mathbf{v} + \mathbf{a}^2\mathbf{a}_2 \cdot \mathbf{v} + \mathbf{a}^3\mathbf{a}_3 \cdot \mathbf{v},$$
$$\mathbf{v} = \mathbf{v} \cdot \mathbf{a}^1\mathbf{a}_1 + \mathbf{v} \cdot \mathbf{a}^2\mathbf{a}_2 + \mathbf{v} \cdot \mathbf{a}^3\mathbf{a}_3;$$

hence:

$$(3) \qquad \mathbf{v} = \mathbf{I} \cdot \mathbf{v}; \qquad \mathbf{v} = \mathbf{v} \cdot \mathbf{I}.$$

A special form of idemfactor is obtained by taking $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ equal to $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, respectively, whereupon, since the reciprocal system to the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system is this system itself, $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ will also be equal to $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, respectively; consequently:

$$\mathrm{I} = \mathbf{ii} + \mathbf{jj} + \mathbf{kk}. \tag{4}$$

The idemfactor is a very special case of a symmetric dyadic. It is a complete dyadic, however, since it is incapable of reduction to a sum of less than three dyads.

§63

The Direct Product of Dyadics

We consider first the products of two dyads, $\mathbf{pq}$ and $\mathbf{rs}$. By definition, for the Direct Products of the two dyads we have:

$$(\mathbf{pq}) \cdot (\mathbf{rs}) = \mathbf{pq} \cdot \mathbf{rs} = \mathbf{q} \cdot \mathbf{r}\mathbf{ps}$$
$$(\mathbf{rs}) \cdot (\mathbf{rq}) = \mathbf{rs} \cdot \mathbf{pq} = \mathbf{s} \cdot \mathbf{p}\mathbf{rq}$$

Accordingly, if in the indicated product of two dyads the two end vectors be called the extremes, and the other two vectors the means, the product is equivalent to a dyad whose antecedent and consequent are the extremes, taken in order, with a scalar coefficient consisting of the scalar product of the means.

The Direct Product of a dyadic Φ into a dyadic Ψ is denoted by $\Phi \cdot \Psi$, and is defined as the dyadic obtained by forming the sum of the direct products of each dyad of Φ into each dyad of Ψ. For example, if:

$$\Phi = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3,$$
$$\Psi = \mathbf{c}_1\mathbf{d}_1 + \mathbf{c}_2\mathbf{d}_2 + \mathbf{c}_3\mathbf{d}_3,$$

then:

$$\begin{aligned} \Phi \cdot \Psi &= (\mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3) \cdot (\mathbf{c}_1\mathbf{d}_1 + \mathbf{c}_2\mathbf{d}_2 + \mathbf{c}_3\mathbf{d}_3) \\ &= \mathbf{b}_1 \cdot \mathbf{c}_1\mathbf{a}_1\mathbf{d}_1 + \mathbf{b}_1 \cdot \mathbf{c}_2\mathbf{a}_1\mathbf{d}_2 + \mathbf{b}_1 \cdot \mathbf{c}_3\mathbf{a}_1\mathbf{d}_3 \\ &\quad + \mathbf{b}_2 \cdot \mathbf{c}_1\mathbf{a}_2\mathbf{d}_1 + \mathbf{b}_2 \cdot \mathbf{c}_2\mathbf{a}_2\mathbf{d}_2 + \mathbf{b}_2 \cdot \mathbf{c}_3\mathbf{a}_2\mathbf{d}_3 \\ &\quad + \mathbf{b}_3 \cdot \mathbf{c}_1\mathbf{a}_3\mathbf{d}_1 + \mathbf{b}_3 \cdot \mathbf{c}_2\mathbf{a}_3\mathbf{d}_2 + \mathbf{b}_3 \cdot \mathbf{c}_3\mathbf{a}_3\mathbf{d}_3. \end{aligned} \tag{1}$$

It is obviously true in general that:

In the expansion of the indicated direct product of two dyadics the distributive law is valid, provided the order of the dyads be maintained in the multiplication of each pair of their respective dyads.

The direct product of a dyadic Φ and an idemfactor I *is equal to the dyadic itself.*

For, if $\mathbf{\Phi}$ be expressed in the form:

$$\mathbf{\Phi} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3,$$

where $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ are non-coplanar, and $\mathbf{I}$ in the form:

$$\mathbf{I} = \mathbf{b}^1\mathbf{b}_1 + \mathbf{b}^2\mathbf{b}_2 + \mathbf{b}^3\mathbf{b}_3;$$

then:

$$\begin{aligned}\mathbf{\Phi}\cdot\mathbf{I} &= (\mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3)\cdot(\mathbf{b}^1\mathbf{b}_1 + \mathbf{b}^2\mathbf{b}_2 + \mathbf{b}^3\mathbf{b}_3),\\ &= \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \mathbf{a}_3\mathbf{b}_3;\end{aligned}$$

hence:

$$\mathbf{\Phi}\cdot\mathbf{I} = \mathbf{\Phi}; \tag{3}$$

in like manner it can be shown that:

$$\mathbf{I}\cdot\mathbf{\Phi} = \mathbf{\Phi}. \tag{4}$$

The direct product of two dyadics and a vector is subject to the associative law, provided the vector follows or precedes both dyadics:

$$\begin{aligned}(\mathbf{\Phi}\cdot\mathbf{\Psi})\cdot\mathbf{v} &= \mathbf{\Phi}\cdot(\mathbf{\Psi}\cdot\mathbf{v}),\\ \mathbf{v}\cdot(\mathbf{\Phi}\cdot\mathbf{\Psi}) &= (\mathbf{v}\cdot\mathbf{\Phi})\cdot\mathbf{\Psi}.\end{aligned}$$

To prove this, it is only necessary to show that corresponding typical terms on the left and right of each of these equations are equal. Let $\mathbf{pq}$ and $\mathbf{rs}$ represent typical dyads of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ respectively. Then:

$$\begin{aligned}(\mathbf{q}\cdot\mathbf{rps})\cdot\mathbf{v} &= \mathbf{pq}\cdot\mathbf{rs}\cdot\mathbf{v} = \mathbf{pq}\cdot(\mathbf{rs}\cdot\mathbf{v}),\\ \mathbf{v}\cdot(\mathbf{q}\cdot\mathbf{rps}) &= \mathbf{v}\cdot\mathbf{pq}\cdot\mathbf{rs} = (\mathbf{v}\cdot\mathbf{pq})\cdot\mathbf{rs}.\end{aligned}$$

These equations prove the equality of the typical terms in question. We can therefore write without ambiguity:

$$(\mathbf{\Phi}\cdot\mathbf{\Psi})\cdot\mathbf{v} = \mathbf{\Phi}\cdot(\mathbf{\Psi}\cdot\mathbf{v}) = \mathbf{\Phi}\cdot\mathbf{\Psi}\cdot\mathbf{v}, \tag{5}$$

$$\mathbf{v}\cdot(\mathbf{\Phi}\cdot\mathbf{\Psi}) = (\mathbf{v}\cdot\mathbf{\Phi})\cdot\mathbf{\Psi} = \mathbf{v}\cdot\mathbf{\Phi}\cdot\mathbf{\Psi}. \tag{6}$$

By an entirely similar argument the truth of the following generalization can be established:

The associative law is valid for the continued product formed in the dot multiplication by a vector, at either or both ends, of the direct product of any number of dyadics.

For example:

$$\begin{aligned}\mathbf{v}\cdot\mathbf{\Phi}\cdot\mathbf{\Psi}\cdot\mathbf{u} &= (\mathbf{v}\cdot\mathbf{\Phi})\cdot(\mathbf{\Psi}\cdot\mathbf{u}) = \mathbf{v}\cdot(\mathbf{\Phi}\cdot\mathbf{\Psi})\cdot\mathbf{u}\\ &= (\mathbf{v}\cdot\mathbf{\Phi}\cdot\mathbf{\Psi})\cdot\mathbf{u} = \mathbf{v}\cdot(\mathbf{\Phi}\cdot\mathbf{\Psi}\cdot\mathbf{u}).\end{aligned} \tag{7}$$

In the case of a continued dot product of dyadics and one or two vectors in which a vector appears in other than an end position, the associative law is not valid in general.

For example:

(8) $$(\mathbf{\Phi} \cdot \mathbf{v}) \cdot \mathbf{\Psi} \neq \mathbf{\Phi} \cdot (\mathbf{v} \cdot \mathbf{\Psi}),$$

in general.

§64
Skew Products of a Dyadic and a Vector

The Skew Products of a dyadic:

$$\mathbf{\Phi} = \mathbf{A}_1\mathbf{B}_1 + \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{A}_n\mathbf{B}_n,$$

and a vector $\mathbf{v}$ are denoted by $\mathbf{\Phi} \times \mathbf{v}$ and by $\mathbf{v} \times \mathbf{\Phi}$. The definitions of these products are quite analogous to the definitions of the direct products of a dyadic and a vector given in Art. 57. Accordingly, we have:

$$\mathbf{\Phi} \times \mathbf{v} = \mathbf{A}_1\mathbf{B}_1 \times \mathbf{v} + \mathbf{A}_2\mathbf{B}_2 \times \mathbf{v} + \cdots + \mathbf{A}_n\mathbf{B}_n \times \mathbf{v};$$
$$\mathbf{v} \times \mathbf{\Phi} = \mathbf{v} \times \mathbf{A}_1\mathbf{B}_1 + \mathbf{v} \times \mathbf{A}_2\mathbf{B}_2 + \cdots + \mathbf{v} \times \mathbf{A}_n\mathbf{B}_n.$$

Evidently, each of these expressions represents a new dyadic.

By the method of comparison of typical terms, exemplified in the preceding article, it can be shown that:

The associative law is valid for the continued product formed in the dot or cross multiplication by a vector, at either or both ends, of the direct product of any number of dyadics.

For example:

$$(\mathbf{v} \times \mathbf{\Phi}) \cdot \mathbf{\Psi} = \mathbf{v} \times (\mathbf{\Phi} \cdot \mathbf{\Psi}) = \mathbf{v} \times \mathbf{\Phi} \cdot \mathbf{\Psi},$$
$$(\mathbf{\Phi} \cdot \mathbf{\Psi}) \times \mathbf{v} = \mathbf{\Phi} \cdot (\mathbf{\Psi} \times \mathbf{v}) = \mathbf{\Phi} \cdot \mathbf{\Psi} \times \mathbf{v},$$
$$(\mathbf{v} \times \mathbf{\Phi}) \cdot \mathbf{u} = \mathbf{v} \times (\mathbf{\Phi} \cdot \mathbf{u}) = \mathbf{v} \times \mathbf{\Phi} \cdot \mathbf{u},$$
$$\mathbf{v} \cdot (\mathbf{\Phi} \times \mathbf{u}) = (\mathbf{v} \cdot \mathbf{\Phi}) \times \mathbf{u} = \mathbf{v} \cdot \mathbf{\Phi} \times \mathbf{u},$$
$$\mathbf{v} \times (\mathbf{\Phi} \times \mathbf{u}) = (\mathbf{v} \times \mathbf{\Phi}) \times \mathbf{u} = \mathbf{v} \times \mathbf{\Phi} \times \mathbf{u}.$$

In the case of a continued product of dyadics and one or two vectors in which a vector, preceded or followed by dot or cross, appears in other than an end position, the associative law is not valid in general.

For example:

$$(\mathbf{\Phi} \times \mathbf{v}) \cdot \mathbf{\Psi} \neq \mathbf{\Phi} \times (\mathbf{v} \cdot \mathbf{\Psi})$$

in general.

Remembering that dot and cross can be interchanged in a scalar triple product of three vectors, the validity of the following useful relations is easily shown:

$$\mathbf{v}\cdot\mathbf{u}\times\mathbf{\Phi} = (\mathbf{v}\times\mathbf{u})\cdot\mathbf{\Phi},$$
$$\mathbf{\Phi}\times\mathbf{v}\cdot\mathbf{u} = \mathbf{\Phi}\cdot(\mathbf{v}\times\mathbf{u}),$$
$$\mathbf{\Phi}\cdot(\mathbf{v}\times\mathbf{\Psi}) = (\mathbf{\Phi}\times\mathbf{v})\cdot\mathbf{\Psi}.$$

§65
Reciprocal Dyadics

Two complete dyadics $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are called Reciprocal Dyadics if their direct products are equal to the idemfactor I; that is, if:[1]

$$\mathbf{\Phi}\cdot\mathbf{\Psi} = \mathbf{I}; \tag{1}$$
$$\mathbf{\Psi}\cdot\mathbf{\Phi} = \mathbf{I}. \tag{2}$$

If one of these equations is true, so also is the other. For, let us assume:

$$\mathbf{\Phi}\cdot\mathbf{\Psi} = \mathbf{I};$$

then, if v represent an arbitrary vector, we have:

$$\mathbf{v}\cdot(\mathbf{\Phi}\cdot\mathbf{\Psi})\cdot\mathbf{\Phi} = \mathbf{v}\cdot\mathbf{I}\cdot\mathbf{\Phi} = \mathbf{v}\cdot\mathbf{\Phi} = (\mathbf{v}\cdot\mathbf{\Phi})\cdot\mathbf{I};$$

but:

$$\mathbf{v}\cdot(\mathbf{\Phi}\cdot\mathbf{\Psi})\cdot\mathbf{\Phi} = (\mathbf{v}\cdot\mathbf{\Phi})\cdot(\mathbf{\Psi}\cdot\mathbf{\Phi});$$

consequently:

$$(\mathbf{v}\cdot\mathbf{\Phi})\cdot(\mathbf{\Psi}\cdot\mathbf{\Phi}) = (\mathbf{v}\cdot\mathbf{\Phi})\cdot\mathbf{I};$$

since v and therefore $\mathbf{v}\cdot\mathbf{\Phi}$ may take on all values, it follows that:

$$\mathbf{\Psi}\cdot\mathbf{\Phi} = \mathbf{I}.$$

To denote the reciprocal of any dyadic $\mathbf{\Phi}$ the symbol $\mathbf{\Phi}^{-1}$ is used.

With the aid of the criterion of equality for two dyadics it can easily be shown that the reciprocals of equal dyadics must be equal.

Let a complete dyadic $\mathbf{\Phi}$ be given in nonian form:

$$\begin{aligned}\mathbf{\Phi} = \;& a_{11}\mathbf{ii} + a_{12}\mathbf{ij} + a_{13}\mathbf{ik}\\ &+ a_{21}\mathbf{ji} + a_{22}\mathbf{jj} + a_{23}\mathbf{jk}\\ &+ a_{31}\mathbf{ki} + a_{32}\mathbf{kj} + a_{33}\mathbf{kk}.\end{aligned} \tag{3}$$

The corresponding nonian form for the reciprocal of $\mathbf{\Phi}$ can be found as follows:

[1] The dyadics $\mathbf{\Phi}$ and $\mathbf{\Psi}$ must be complete since I is a complete dyadic, and the direct product of two dyadics either of which is not complete cannot give a complete dyadic.

Let $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ denote the vectors[1)] produced by $\boldsymbol{\Phi}$ acting as a prefactor upon $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ respectively, then $\boldsymbol{\Phi}$ can also be written in the form:

$$\boldsymbol{\Phi} = \mathbf{b}_1\mathbf{i} + \mathbf{b}_2\mathbf{j} + \mathbf{b}_3\mathbf{k}, \tag{4}$$

in accordance with the criteria for equality of two dyadics; and, if $\mathbf{b}^1$, $\mathbf{b}^2$, $\mathbf{b}^3$ are vectors of the system reciprocal to the $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$-system, then:

$$\boldsymbol{\Phi}^{-1} = \mathbf{i}\mathbf{b}^1 + \mathbf{j}\mathbf{b}^2 + \mathbf{k}\mathbf{b}^3, \tag{5}$$

since this dyadic acting as a prefactor upon $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, produces $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. Upon comparison of the forms (3) and (4) we find:

$$\begin{aligned} \mathbf{b}_1 &= a_{11}\mathbf{i} + a_{21}\mathbf{j} + a_{31}\mathbf{k}, \\ \mathbf{b}_2 &= a_{12}\mathbf{i} + a_{22}\mathbf{j} + a_{32}\mathbf{k}, \\ \mathbf{b}_3 &= a_{13}\mathbf{i} + a_{23}\mathbf{j} + a_{33}\mathbf{k}. \end{aligned}$$

Solving these linear equations for $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, we get:

$$\begin{aligned} & \frac{A_{11}}{\Delta_a}\mathbf{b}_1 + \frac{A_{12}}{\Delta_a}\mathbf{b}_2 + \frac{A_{13}}{\Delta_a}\mathbf{b}_3, \\ \mathbf{j} \quad & \frac{A_{21}}{\Delta_a}\mathbf{b}_1 + \frac{A_{22}}{\Delta_a}\mathbf{b}_2 + \frac{A_{23}}{\Delta_a}\mathbf{b}_3, \\ \mathbf{k} = & \frac{A_{31}}{\Delta_a}\mathbf{b}_1 + \frac{A_{32}}{\Delta_a}\mathbf{b}_2 + \frac{A_{33}}{\Delta_a}\mathbf{b}_3, \end{aligned} \tag{6}$$

where:

$$\Delta_a = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{7}$$

and A_{rs} is the cofactor of the typical term a_{rs} in this determinant, defined as the first minor obtained by suppressing the r'th row and the s'th column with + or − sign, according as $r + s$ is even or odd. Upon scalar multiplication of each of equations (6) by $\mathbf{b}^1$, $\mathbf{b}^2$, $\mathbf{b}^3$ in turn we find:

$$\begin{aligned} \mathbf{b}^1 \cdot \mathbf{i} &= \frac{A_{11}}{\Delta_a}, & \mathbf{b}^1 \cdot \mathbf{j} &= \frac{A_{21}}{\Delta_a}, & \mathbf{b}^1 \cdot \mathbf{k} &= \frac{A_{31}}{\Delta_a}, \\ \mathbf{b}^2 \cdot \mathbf{i} &= \frac{A_{12}}{\Delta_a}, & \mathbf{b}^2 \cdot \mathbf{j} &= \frac{A_{22}}{\Delta_a}, & \mathbf{b}^2 \cdot \mathbf{k} &= \frac{A_{32}}{\Delta_a}, \\ \mathbf{b}^3 \cdot \mathbf{i} &= \frac{A_{13}}{\Delta_a}, & \mathbf{b}^3 \cdot \mathbf{j} &= \frac{A_{23}}{\Delta_a}, & \mathbf{b}^3 \cdot \mathbf{k} &= \frac{A_{33}}{\Delta_a}. \end{aligned}$$

[1)] These vectors must be non-coplanar, since $\boldsymbol{\Phi}$ is assumed to be a complete dyadic.

These equations give the measure-numbers of the i, j, k-components of the vectors $\mathbf{b}^1$, $\mathbf{b}^2$, $\mathbf{b}^3$, and for these vectors we can therefore write:

$$
\begin{aligned}
\mathbf{b}^1 &= \frac{A_{11}}{\Delta_a}\mathbf{i} + \frac{A_{21}}{\Delta_a}\mathbf{j} + \frac{A_{31}}{\Delta_a}\mathbf{k}, \\
\mathbf{b}^2 &= \frac{A_{12}}{\Delta_a}\mathbf{i} + \frac{A_{22}}{\Delta_a}\mathbf{j} + \frac{A_{32}}{\Delta_a}\mathbf{k}, \qquad (8) \\
\mathbf{b}^3 &= \frac{A_{13}}{\Delta_a}\mathbf{i} + \frac{A_{23}}{\Delta_a}\mathbf{j} + \frac{A_{33}}{\Delta_a}\mathbf{k}.
\end{aligned}
$$

Upon substitution of these expressions for $\mathbf{b}^1$, $\mathbf{b}^2$, $\mathbf{b}^3$ in equation (5) we find for the nonian form of the reciprocal of $\mathbf{\Phi}$:

$$
\begin{aligned}
\mathbf{\Phi}^{-1} = \frac{1}{\Delta_a}[&A_{11}\mathbf{ii} + A_{21}\mathbf{ij} + A_{31}\mathbf{ik} \\
&+ A_{12}\mathbf{ji} + A_{22}\mathbf{jj} + A_{32}\mathbf{jk} \qquad (9) \\
&+ A_{13}\mathbf{ki} + A_{23}\mathbf{kj} + A_{33}\mathbf{kk}].
\end{aligned}
$$

The dyadic enclosed in brackets is called the Adjunct of $\mathbf{\Phi}$; it will be denoted by $\mathbf{\Phi}_{adj}$, and its determinant by Δ_A, so that:

$$
\begin{aligned}
\mathbf{\Phi}_{adj} = \Delta_a\mathbf{\Phi}^{-1} = &A_{11}\mathbf{ii} + A_{21}\mathbf{ij} + A_{31}\mathbf{ik} \\
&+ A_{12}\mathbf{ji} + A_{22}\mathbf{jj} + A_{32}\mathbf{jk} \qquad (10) \\
&+ A_{13}\mathbf{ki} + A_{23}\mathbf{kj} + A_{33}\mathbf{kk};
\end{aligned}
$$

$$
\Delta_A = \begin{vmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{vmatrix} \qquad (11)
$$

The elements of Δ_A are the cofactors of the elements of Δ_a.

Using equations (8), it can easily be shown that:

$$\mathbf{b}^1 \cdot \mathbf{b}^2 \times \mathbf{b}^3 = \frac{\Delta_A}{\Delta_a^3};$$

and since, by equation (5), Art. 19:

$$\mathbf{b}^1 \cdot \mathbf{b}^2 \times \mathbf{b}^3 = \frac{1}{[\mathbf{b}_1\mathbf{b}_2\mathbf{b}_3]} = \frac{1}{\Delta_a},$$

it follows that:

$$\Delta_A = \Delta_a^2. \qquad (12)$$

From equations (1), (9), and (10) the following relation, which will be of use later, follows directly:

$$\mathbf{\Phi} \cdot \mathbf{\Phi}_{adj} = \Delta_a\mathbf{I}. \qquad (13)$$

§66

Several Important Theorems Relating to Dyadic Operators

In working with dyadic operators, the following theorems are often found useful:

(a) *The conjugate of the sum (difference) of two dyadics is equal to the sum (difference) of the conjugates of the individual dyadics.*

Let $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denote the dyadics. Then:

(1) $$(\mathbf{\Phi} + \mathbf{\Psi})_c = \mathbf{\Phi}_c + \mathbf{\Psi}_c, \qquad (\mathbf{\Phi} - \mathbf{\Psi})_c = \mathbf{\Phi}_c - \mathbf{\Psi}_c.$$

The truth of this theorem is directly evident from the definitions of a dyadic and its conjugate given in Art. 57.

(b) *The conjugate of the direct product of any number of dyadics is equal to the direct product of their individual conjugates taken in reverse order.*

It will suffice to give the proof of this theorem for the case of two dyadics, since that for the general case can then be directly inferred.

Let, then, $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denote the two dyadics. We wish to prove that:

(2) $$(\mathbf{\Phi} \cdot \mathbf{\Psi})_c = \mathbf{\Psi}_c \cdot \mathbf{\Phi}_c.$$

We have:

$$\begin{aligned}(\mathbf{\Phi} \cdot \mathbf{\Psi})_c \cdot \mathbf{v} &= \mathbf{v} \cdot (\mathbf{\Phi} \cdot \mathbf{\Psi}) = (\mathbf{v} \cdot \mathbf{\Phi}) \cdot \mathbf{\Psi} = (\mathbf{\Phi}_c \cdot \mathbf{v}) \cdot \mathbf{\Psi} \\ &= \mathbf{\Psi}_c \cdot (\mathbf{\Phi}_c \cdot \mathbf{v}) = (\mathbf{\Psi}_c \cdot \mathbf{\Phi}_c) \cdot \mathbf{v}, \text{ for all values of } \mathbf{v};\end{aligned}$$

hence, equation (2) is valid.

(c) *The sum (difference) of a dyadic and its conjugate is a symmetric (anti-symmetric) dyadic.*

If $\mathbf{\Phi}$ denote the dyadic, we have, with the aid of theorem (a):

(3) $$(\mathbf{\Phi} + \mathbf{\Phi}_c)_c = \mathbf{\Phi}_c + (\mathbf{\Phi}_c)_c = \mathbf{\Phi} + \mathbf{\Phi}_c,$$

(4) $$(\mathbf{\Phi} - \mathbf{\Phi}_c)_c = \mathbf{\Phi}_c - (\mathbf{\Phi}_c)_c = -(\mathbf{\Phi} - \mathbf{\Phi}_c).$$

These equations suffice to prove the theorem.

(d) *Any dyadic can be resolved into a symmetric and an anti-symmetric part.*[1]

Let $\mathbf{\Phi}$ denote the dyadic. Then:

(5) $$\mathbf{\Phi} = \left(\frac{\mathbf{\Phi} + \mathbf{\Phi}_c}{2}\right) + \left(\frac{\mathbf{\Phi} - \mathbf{\Phi}_c}{2}\right), \text{ identically.}$$

[1] This can be done in but one way. Cf. Gibbs-Wilson, Vector Analysis, p. 296.

In accordance with theorem (c) the first (second) bracketed expression on the right is a symmetric (anti-symmetric) dyadic.

(e) *The conjugate of the reciprocal of a complete dyadic is equal to the reciprocal of its conjugate.*

Let $\mathbf{\Phi}$ denote the dyadic. Then, with the aid of theorem (b):

$$(\mathbf{\Phi}^{-1})_c \cdot \mathbf{\Phi}_c = (\mathbf{\Phi} \cdot \mathbf{\Phi}^{-1})_c;$$

but, in accordance with the definition of a reciprocal dyadic:

$$(\mathbf{\Phi} \cdot \mathbf{\Phi}^{-1})_c = \mathbf{I}_c, \qquad (\mathbf{\Phi}_c)^{-1} \cdot \mathbf{\Phi}_c = \mathbf{I};$$

hence, noting that $\mathbf{I}_c = \mathbf{I}$:

$$(\mathbf{\Phi}_c)^{-1} \cdot \mathbf{\Phi}_c = (\mathbf{\Phi}^{-1})_c \cdot \mathbf{\Phi}_c;$$

consequently:

$$(\mathbf{\Phi}_c)^{-1} \cdot (\mathbf{\Phi}_c \cdot \mathbf{v}) = (\mathbf{\Phi}^{-1})_c \cdot (\mathbf{\Phi}_c \cdot \mathbf{v}),$$

for all values of $\mathbf{v}$, and therefore of $\mathbf{\Phi}_c \cdot \mathbf{v}$, hence:

(6) $$(\mathbf{\Phi}_c)^{-1} = (\mathbf{\Phi}^{-1})_c.$$

The symbol $\mathbf{\Phi}_c^{-1}$ can, therefore, be used without ambiguity to denote the reciprocal of the conjugate of $\mathbf{\Phi}$ or the conjugate of the reciprocal of $\mathbf{\Phi}$.

(f) *The reciprocal of the direct product of two complete dyadics is equal to the product of their individual reciprocals taken in reverse order.*

If $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denote the two dyadics, then:

$$(\mathbf{\Phi} \cdot \mathbf{\Psi}) \cdot (\mathbf{\Psi}^{-1} \cdot \mathbf{\Phi}^{-1}) = \mathbf{\Phi} \cdot (\mathbf{\Psi} \cdot \mathbf{\Psi}^{-1}) \cdot \mathbf{\Phi}^{-1} = \mathbf{\Phi} \cdot \mathbf{I} \cdot \mathbf{\Phi}^{-1} = \mathbf{\Phi} \cdot \mathbf{\Phi}^{-1} = \mathbf{I}.$$

Therefore, $\mathbf{\Phi} \cdot \mathbf{\Psi}$ and $\mathbf{\Psi}^{-1} \cdot \mathbf{\Phi}^{-1}$ must be reciprocal dyadics, and hence:

$$(\mathbf{\Phi} \cdot \mathbf{\Psi})^{-1} = \mathbf{\Psi}^{-1} \cdot \mathbf{\Phi}^{-1}.$$

By induction the theorem just stated can be generalized so as to read as follows:

(g) *The reciprocal of the direct product of any number of complete dyadics is equal to the direct product of their individual reciprocals taken in reverse order.*

§67

Reduction of a Dyadic to a Normal Form

Any complete dyadic can be reduced to a trinomial form in which the vectors of the antecedents are mutually perpendicular, and likewise those of the consequents. In fact, any complete dyadic $\mathbf{\Phi}$ can be reduced to the form:

(1) $$\mathbf{\Phi} = A\mathbf{i}'\mathbf{i} + B\mathbf{j}'\mathbf{j} + C\mathbf{k}'\mathbf{k},$$

where $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are congruent, right-handed systems of unit vectors, and A, B, C are all positive or all negative numbers.

In proving this we shall make use of properties connected with the affine transformation of a sphere of unit radius.

If $\mathbf{r}$ denote the position-vector of any point P on the sphere with respect to its center as origin, and $\mathbf{q}$ the linear vector function produced by $\boldsymbol{\Phi}$ acting as prefactor upon $\mathbf{r}$, we can write:

$$\mathbf{r} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3, \tag{2}$$

$$\mathbf{q} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3, \tag{3}$$

where $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are a set of special values of $\mathbf{r}$ constituting an *orthogonal* base-system of unit vectors, x_1, x_2, x_3 are the co-ordinates of P on this system, and where:

$$\mathbf{b}_1 = \boldsymbol{\Phi}\cdot(\mathbf{a}_1), \qquad \mathbf{b}_2 = \boldsymbol{\Phi}\cdot(\mathbf{a}_2), \qquad \mathbf{b}_3 = \boldsymbol{\Phi}\cdot(\mathbf{a}_3). \tag{4}$$

The linear vector function $\mathbf{q}$ will be non-degenerate, since $\boldsymbol{\Phi}$ is supposed a complete dyadic. Hence, $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ will be non-coplanar vectors, and may therefore be taken as a base-system of reference for the vector $\mathbf{q}$ which we shall consider as the position-vector of a point Q. On this system x_1, x_2, x_3 are co-ordinates of the point Q.

From equation (2), upon forming the scalar product $\mathbf{r}\cdot\mathbf{r}$, we find:

$$x_1^2 + x_2^2 + x_3^2 = 1. \tag{5}$$

On the $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$-base-system this equation is that of the unit sphere. On the $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$-base-system it is the equation of the surface which is the locus of the point Q. Since this surface must be a closed surface, and since its equation is of the second degree, it must be an ellipsoid. It will be called the Tensor Ellipsoid. We shall first show that the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ must coincide in direction with conjugate diameters of the ellipsoid, the magnitudes of these vectors being equal to those of the corresponding semi-conjugate diameters.

Upon differentiation of equations (3) and (5) we get:

$$d\mathbf{q} = dx_1\mathbf{b}_1 + dx_2\mathbf{b}_2 + dx_3\mathbf{b}_3; \tag{6}$$

$$0 = x_1dx_1 + x_2dx_2 + x_3ax_3. \tag{7}$$

In the first of these equations $d\mathbf{q}$ represents the increment in the position vector $\mathbf{q}$ of any point $Q(x_1, x_2, x_3)$ on the ellipsoid and must of course be parallel to the tangent plane at Q. Now, when $\mathbf{q} = \mathbf{b}_1$, it follows by equation (3) that: $x_1 = 1, x_2 = x_3 = 0$; and hence by equation (7) that: $dx_1 = 0$. It then follows from equation (6) that

when $\mathbf{q} = \mathbf{b}_1$ the tangent plane must be parallel to the diametral plane determined by $\mathbf{b}_2$, $\mathbf{b}_3$. In like manner it can be shown that when $\mathbf{q} = \mathbf{b}_2$, $\mathbf{q} = \mathbf{b}_3$ the corresponding tangent planes must be parallel to the diametral planes determined by $\mathbf{b}_3$, $\mathbf{b}_1$ and by $\mathbf{b}_1$, $\mathbf{b}_2$, respectively. Hence, $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ specify semi-conjugate diameters of the ellipsoid. Now suppose $\mathbf{a}_1$ so chosen that $\mathbf{b}_1$ $(= \mathbf{\Phi} \cdot \mathbf{a}_1)$ specifies a semi-principal axis; then $\mathbf{b}_2$ and $\mathbf{b}_3$ must be perpendicular to $\mathbf{b}_1$. Next, suppose $\mathbf{a}_2$ so chosen that $\mathbf{b}_2$ $(= \mathbf{\Phi} \cdot \mathbf{a}_2)$ specifies a second semi-principal axis; then $\mathbf{b}_3$ must specify the third semi-principal axis. Hence, there must be a set of values for $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, say $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, which, when acted upon by the dyadic $\mathbf{\Phi}$ as a prefactor, transform into an orthogonal set of vectors specifying semi-principal axes of the tensor ellipsoid. If, therefore, $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ constitute a right-handed orthogonal set of unit vectors collinear in direction with these axes, respectively, the dyadic $\mathbf{\Phi}$ can be put in the form:

$$\mathbf{\Phi} = A\mathbf{i}'\mathbf{i} + B\mathbf{j}'\mathbf{j} + C\mathbf{k}'\mathbf{k},$$

where A, B, C are positive or negative scalars whose magnitudes are, respectively, those of the semi-principal axes of the ellipsoid. Now the signs of any two of the coefficients A, B, C can be reversed without disturbing the right-handed relationship of the vectors $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ by simply reversing the direction of any two of these vectors. It follows that $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ can be so chosen that the coefficients A, B, C are either all positive or all negative. The proposition stated in the first paragraph of the present article is thus established.

In the demonstration just given $\mathbf{\Phi}$ was assumed to be a complete dyadic. If $\mathbf{\Phi}$ is assumed to be a planar dyadic, a similar demonstration shows that it can be reduced to a normal form:

$$\mathbf{\Phi} = A\mathbf{i}'\mathbf{i} + B\mathbf{j}'\mathbf{j}, \tag{8}$$

where the coefficients A and B can be taken as positive. Similarly, if $\mathbf{\Phi}$ be assumed to be a linear dyadic, it can be shown to be reducible to the normal form:

$$\mathbf{\Phi} = A\mathbf{i}'\mathbf{i}, \tag{9}$$

where A is a positive coefficient.

If, in the reduction of a complete dyadic $\mathbf{\Phi}$, the coefficients A, B, C are not all different, the reduction can be effected in an infinite number of ways. If two of the coefficients, say B and C, are equal, the tensor ellipsoid reduces to one of revolution about the unique axis $\mathbf{i}'$, and therefore $\mathbf{j}'$ together with $\mathbf{k}'$ can be rotated

about this unique axis without affecting the value of $\mathbf{\Phi}$. If all three of the coefficients A, B, C are equal, the tensor ellipsoid degenerates into a sphere, and in this case the $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$-system can be rotated in any manner without affecting the value of $\mathbf{\Phi}$.

Normal Form for a Symmetric Dyadic

Any symmetric complete dyadic $\mathbf{\Phi}$ can be reduced to the form:

$$\mathbf{\Phi} = a\mathbf{ii} + b\mathbf{jj} + c\mathbf{kk}, \tag{1}$$

where the coefficients a, b, c are positive or negative constants.

The dyadic $\mathbf{\Phi}$ and its conjugate $\mathbf{\Phi}_c$ can, by equation (1), Art. 67. be expressed in the forms:

$$\mathbf{\Phi} = A\mathbf{i}'\mathbf{i} + B\mathbf{j}'\mathbf{j} + C\mathbf{k}'\mathbf{k}, \tag{2}$$

$$\mathbf{\Phi}_c = A\mathbf{i}\mathbf{i}' + B\mathbf{j}\mathbf{j}' + C\mathbf{k}\mathbf{k}'. \tag{3}$$

If $\mathbf{\Phi}$ is symmetric, then:

$$\mathbf{\Phi} \cdot \mathbf{\Phi}_c = \mathbf{\Phi}_c \cdot \mathbf{\Phi}.$$

Hence:

$$A^2\mathbf{i}'\mathbf{i}' + B^2\mathbf{j}'\mathbf{j}' + C^2\mathbf{k}'\mathbf{k}' = A^2\mathbf{ii} + B^2\mathbf{jj} + C^2\mathbf{kk},$$

and since:

$$\mathbf{i}'\mathbf{i}' + \mathbf{j}'\mathbf{j}' + \mathbf{k}'\mathbf{k}' = \mathbf{ii} + \mathbf{jj} + \mathbf{kk},$$

the expressions on the left and right being idemfactors, we shall have, after multiplication of the last equation by A^2 and subtraction from the preceding equation:

$$(B^2 - A^2)\,\mathbf{j}'\mathbf{j}' + (C^2 - A^2)\,\mathbf{k}'\mathbf{k}' = (B^2 - A^2)\,\mathbf{jj} + (C^2 - A^2)\,\mathbf{kk}.$$

The dyadic in the right-hand member of this equation acting as a prefactor or postfactor upon the unit vector $\mathbf{i}$ yields a null-vector, and that in the left-hand member must therefore do likewise; hence:

$$(B^2 - A^2)\,\mathbf{j}' \cdot \mathbf{i}\mathbf{j}' + (C^2 - A^2)\,\mathbf{k}' \cdot \mathbf{i}\mathbf{k}' = 0. \tag{4}$$

If A, B, C are unequal, it is therefore necessary that:

$$\mathbf{j}' \cdot \mathbf{i} = \mathbf{k}' \cdot \mathbf{i} = 0.$$

Hence $\mathbf{j}'$ and $\mathbf{k}'$ must be perpendicular to $\mathbf{i}$ and, since $\mathbf{j}'$ and $\mathbf{k}'$ are also perpendicular to $\mathbf{i}'$, it follows that $\mathbf{i}'$ and $\mathbf{i}$ must be collinear. In a similar way $\mathbf{j}'$, $\mathbf{j}$ and $\mathbf{k}'$, $\mathbf{k}$, respectively, can be shown to be collinear. Hence, in equation (2), for $A\mathbf{i}'$, $B\mathbf{j}'$, $C\mathbf{k}'$ we can put

$a\mathbf{i}$, $b\mathbf{j}$, $c\mathbf{k}$, if a, b, c denote appropriate positive or negative numbers, and thus obtain the form (1) for $\mathbf{\Phi}$.

Suppose now that $C = B \neq A$. This corresponds to the case for which the tensor ellipsoid is one of rotation about the axis parallel to $\mathbf{i}'$. It follows from equation (4) that $\mathbf{i}'$ must be collinear with $\mathbf{i}$; furthermore, since in this case, as noted in the preceding article, the two vectors $\mathbf{j}'$, $\mathbf{k}'$ can be rotated together about $\mathbf{i}'$ as an axis without changing the value of $\mathbf{\Phi}$, they can be so rotated as to become collinear with $\mathbf{j}$, $\mathbf{k}$, respectively. The theorem, therefore, also holds for this case and, it can be shown in like manner to hold for the cases $B = A \neq C$, and $A = C \neq B$.

Suppose next that $A = B = C$. In this case the tensor ellipsoid degenerates into a sphere, and, as pointed out in the preceding article, the $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$-system of axes can be rotated about the origin without changing the value of the dyadic $\mathbf{\Phi}$. It can, therefore, be brought into coincidence with the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system, and the theorem is therefore valid for this case.

§69

Dyadic Invariants

Certain combinations of the nine coefficients in the nonian form of a dyadic have the same values for all $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-systems of axes. These are called Dyadic Invariants. Three of these, which are of fundamental importance, we shall now find by a method which will also furnish results which will be useful in connection with the consideration to follow of certain special forms of dyadics.

Let a dyadic $\mathbf{\Phi}$ be given in the nonian form:

$$\begin{aligned} \mathbf{\Phi} = \;& a_{11}\mathbf{ii} + a_{12}\mathbf{ij} + a_{13}\mathbf{ik} \\ & + a_{21}\mathbf{ji} + a_{22}\mathbf{jj} + a_{23}\mathbf{jk} \\ & + a_{31}\mathbf{ki} + a_{32}\mathbf{kj} + a_{33}\mathbf{kk}. \end{aligned} \tag{1}$$

If we let:

$$\begin{aligned} \mathbf{b}_1 &= a_{11}\mathbf{i} + a_{21}\mathbf{j} + a_{31}\mathbf{k} = \mathbf{\Phi}\cdot\mathbf{i}, \\ \mathbf{b}_2 &= a_{12}\mathbf{i} + a_{22}\mathbf{j} + a_{32}\mathbf{k} = \mathbf{\Phi}\cdot\mathbf{j}, \\ \mathbf{b}_3 &= a_{13}\mathbf{i} + a_{23}\mathbf{j} + a_{33}\mathbf{k} = \mathbf{\Phi}\cdot\mathbf{k}, \end{aligned} \tag{2}$$

then $\mathbf{\Phi}$ can be written in the trinomial form:

$$\mathbf{\Phi} = \mathbf{b}_1\mathbf{i} + \mathbf{b}_2\mathbf{j} + \mathbf{b}_3\mathbf{k}. \tag{3}$$

If we wish to find the directions in space for which any vector $\mathbf{r}$, when operated upon by $\mathbf{\Phi}$ as a prefactor, will be transformed

into a vector proportional to $\mathbf{r}$, the necessary and sufficient condition is expressed by the equation:

$$\mathbf{\Phi}\cdot\mathbf{r} = \lambda\mathbf{r}, \tag{4}$$

where λ is a factor of proportionality. If $\mathbf{r}$ be expressed in terms of its components on the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-base-system as follows:

$$\mathbf{r} = \mathbf{i}\cdot\mathbf{ri} + \mathbf{j}\cdot\mathbf{rj} + \mathbf{k}\cdot\mathbf{rk};$$

then:

$$\begin{aligned}\mathbf{\Phi}\cdot\mathbf{r} &= \mathbf{i}\cdot\mathbf{r\Phi}\cdot\mathbf{i} + \mathbf{j}\cdot\mathbf{r\Phi}\cdot\mathbf{j} + \mathbf{k}\cdot\mathbf{r\Phi}\cdot\mathbf{k}\\ &= \mathbf{i}\cdot\mathbf{rb}_1 + \mathbf{j}\cdot\mathbf{rb}_2 + \mathbf{k}\cdot\mathbf{rb}_3.\end{aligned}$$

The condition (4) can then be expressed in the form:

$$\mathbf{i}\cdot\mathbf{r}(\lambda\mathbf{i} - \mathbf{b}_1) + \mathbf{j}\cdot\mathbf{r}(\lambda\mathbf{j} - \mathbf{b}_2) + \mathbf{k}\cdot\mathbf{r}(\lambda\mathbf{k} - \mathbf{b}_3) = 0.$$

Since the coefficients of the vectors enclosed by brackets cannot all be zero, these vectors are necessarily coplanar. Hence:

$$(\lambda\mathbf{i} - \mathbf{b}_1)\cdot(\lambda\mathbf{j} - \mathbf{b}_2)\times(\lambda\mathbf{k} - \mathbf{b}_3) = 0.$$

This is a cubic equation in λ which, upon expansion of the left-hand member, can be written in the form:

$$\lambda^3 - \Phi_1\lambda^2 + \Phi_2\lambda - \Phi_3 = 0, \tag{5}$$

where:

$$\Phi_1 = \mathbf{b}_1\cdot\mathbf{i} + \mathbf{b}_2\cdot\mathbf{j} + \mathbf{b}_3\cdot\mathbf{k}, \tag{6}$$

$$\Phi_2 = \mathbf{b}_2\times\mathbf{b}_3\cdot\mathbf{i} + \mathbf{b}_3\times\mathbf{b}_1\cdot\mathbf{j} + \mathbf{b}_1\times\mathbf{b}_2\cdot\mathbf{k}, \tag{7}$$

$$\Phi_3 = \mathbf{b}_1\cdot\mathbf{b}_2\times\mathbf{b}_3. \tag{8}$$

Now, since the values of λ must be independent of the manner in which the dyadic $\mathbf{\Phi}$ is expressed, it follows that the roots of this cubic equation in λ, and therefore the values of Φ_1, Φ_2, and Φ_3, must have the same values on all right-handed $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-systems. The quantities Φ_1, Φ_2, Φ_3 are, therefore, called Dyadic Invariants for such systems.

From equations (6), (7), and (8), with the aid of equations (2), we find;

$$\Phi_1 = a_{11} + a_{22} + a_{33}, \tag{6'}$$

$$\Phi_2 = a_{11}a_{22} + a_{22}a_{33} + a_{33}a_{11} - a_{12}a_{21} - a_{23}a_{32} - a_{31}a_{13}, \tag{7'}$$

$$\Phi_3 \quad \begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} \tag{8'}$$

From these expressions it is seen that Φ_1 is equal to the sum of the terms in the principal diagonal of the determinant expressing Φ_3, and that Φ_2 is the sum of the diagonal minors of Φ_3.

If $\mathbf{\Phi}$ is a complete dyadic, Φ_2 can be given another and important form. In this case the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ will have a reciprocal system:

$$\mathbf{b}^1 = \frac{\mathbf{b}_2 \times \mathbf{b}_3}{\Delta_a}, \qquad \mathbf{b}^2 = \frac{\mathbf{b}_3 \times \mathbf{b}_1}{\Delta_a}, \qquad \mathbf{b}^3 = \frac{\mathbf{b}_1 \times \mathbf{b}_2}{\Delta_a}$$

where $\Delta_a(= \mathbf{b}_1 \cdot \mathbf{b}_2 \times \mathbf{b}_3 = \Phi_3)$ is the determinant of $\mathbf{\Phi}$ when expressed in nonian form. These equations, with equation (7), give:

$$\Phi_2 = \Delta_a(\mathbf{b}^1 \cdot \mathbf{i} + \mathbf{b}^2 \cdot \mathbf{j} + \mathbf{b}^3 \cdot \mathbf{k}),$$

or, upon using the expressions given by equations (6), Art. 65, for $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ in terms of $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$:

$$(7'') \qquad \Phi_2 = A_{11} + A_{22} + A_{33}.$$

The expression on the right is the sum of the terms of the principal diagonal of the determinant

$$(9) \qquad \Delta_A = \begin{vmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{vmatrix},$$

which, as shown in Art. 65, is the determinant of the dyadic called the Adjunct of $\mathbf{\Phi}$ which is equal to $\Delta_a \mathbf{\Phi}^{-1}$, the elements of this determinant being cofactors of the elements of the determinant Δ_a of the dyadic $\mathbf{\Phi}$ itself when expressed in nonian form.

§70

The Hamilton-Cayley Equation

The dyadic $\mathbf{\Phi} - \lambda\mathbf{I}$ is called the Characteristic Dyadic of $\mathbf{\Phi}$, and can be written in the nonian form:

$$(1) \qquad \begin{aligned} \mathbf{\Phi} - \lambda\mathbf{I} = (a_{11} - \lambda)\mathbf{ii} + a_{12}\mathbf{ij} + a_{13}\mathbf{ik} \\ + a_{21}\mathbf{ji} + (a_{22} - \lambda)\mathbf{jj} + a_{23}\mathbf{jk} \\ + a_{31}\mathbf{ki} + a_{32}\mathbf{kj} + (a_{33} - \lambda)\mathbf{kk}. \end{aligned}$$

The Characteristic Determinant of $\mathbf{\Phi}$, denoted by $\phi(\lambda)$, is the determinant of $\mathbf{\Phi} - \lambda\mathbf{I}$, so that:

$$(2) \qquad \phi(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix}$$

This function is a cubic in λ and, as is easily seen upon expanding and taking account of equations (6′), (7′), and (8′), Art. 69:

$$\phi(\lambda) = -\lambda^3 + \Phi_1\lambda^2 - \Phi_2\lambda + \Phi_3. \tag{3}$$

As seen above, the roots of the scalar equation:

$$\phi(\lambda) = -\lambda^3 + \Phi_1\lambda^2 - \Phi_2\lambda + \Phi_3 = 0 \tag{4}$$

determine the directions in space for which $\mathbf{\Phi} \cdot \mathbf{r} = \lambda\mathbf{r}$.

We shall now show that the equation obtained upon the substitution of $\mathbf{\Phi}$ for λ and $\Phi_3\mathbf{I}$ for Φ_3 in equation (4), viz:

$$\mathbf{\Phi}^3 - \Phi_1\mathbf{\Phi}^2 + \Phi_2\mathbf{\Phi} - \Phi_3\mathbf{I} = 0,$$

is identically true, with the understanding that $\mathbf{\Phi}^2 \equiv \mathbf{\Phi} \cdot \mathbf{\Phi}$ and $\mathbf{\Phi}^3 \equiv \mathbf{\Phi} \cdot \mathbf{\Phi}^2$. By equation (13) Art. 65, noting that $\Delta_a = \Phi_3$:

$$\mathbf{\Phi} \cdot \mathbf{\Phi}_{\text{adj}} = \Phi_3\mathbf{I}.$$

Hence, upon writing $\mathbf{\Phi} - \lambda\mathbf{I}$ for $\mathbf{\Phi}$, we shall have:

$$(\mathbf{\Phi} - \lambda\mathbf{I}) \cdot (\mathbf{\Phi} - \lambda\mathbf{I})_{\text{adj}} = \phi(\lambda)\mathbf{I},$$

or:

$$\mathbf{\Phi} \cdot (\mathbf{\Phi} - \lambda\mathbf{I})_{\text{adj}} - \lambda(\mathbf{\Phi} - \lambda\mathbf{I})_{\text{adj}} = (-\lambda^3 + \Phi_1\lambda^2 - \Phi_2\lambda + \Phi_3)\,\mathbf{I}. \tag{5}$$

The adjunct of $\mathbf{\Phi} - \lambda\mathbf{I}$ can be evaluated in the same manner as the adjunct of $\mathbf{\Phi}$ in Art. 65. We thus find:

$$\begin{aligned}(\mathbf{\Phi} - \lambda\mathbf{I})_{\text{adj}} = {} & C_{11}\mathbf{ii} + C_{21}\mathbf{ij} + C_{31}\mathbf{ik} \\ & + C_{12}\mathbf{ji} + C_{22}\mathbf{jj} + C_{32}\mathbf{jk} \\ & + C_{13}\mathbf{ki} + C_{23}\mathbf{kj} + C_{33}\mathbf{kk},\end{aligned} \tag{6}$$

where:

$$C_{rs} = (-1)^{r+s}\,\phi(\lambda)_{rs}, \tag{7}$$

and $\phi(\lambda)_{rs}$ represents the determinant of the second order obtained by suppressing the r'th row and the s'th column of the determinant $\phi(\lambda)$ given by equation (2). Evidently, C_{rs} will be of the second or lower degree in λ, and we can, therefore, write:

$$(\mathbf{\Phi} - \lambda\mathbf{I})_{\text{adj}} = \mathbf{\Psi}_2\lambda^2 + \mathbf{\Psi}_1\lambda + \mathbf{\Psi}_0, \tag{8}$$

where the $\mathbf{\Psi}$'s are dyadic coefficients free of λ, which for our purposes it is unnecessary to specify further. From equations (5) and (8):

$$\begin{aligned}\mathbf{\Phi} \cdot (\mathbf{\Psi}_2\lambda^2 + \mathbf{\Psi}_1\lambda + \mathbf{\Psi}_0) - \lambda(\mathbf{\Psi}_2\lambda^2 + \mathbf{\Psi}_1\lambda + \mathbf{\Psi}_0) \\ = (-\lambda^3 + \Phi_1\lambda^2 - \Phi_2\lambda + \Phi_3)\,\mathbf{I}.\end{aligned}$$

Upon equating the terms independent of λ on the left and right of this equation, and likewise the coefficients of like powers of λ, we obtain:

$$\begin{aligned}\mathbf{\Phi}\cdot\mathbf{\Psi}_0 &= \Phi_3\mathbf{I},\\ \mathbf{\Phi}\cdot\mathbf{\Psi}_1 - \mathbf{\Psi}_0 &= -\Phi_2\mathbf{I},\\ \mathbf{\Phi}\cdot\mathbf{\Psi}_2 - \mathbf{\Psi}_1 &= \Phi_1\mathbf{I},\\ -\mathbf{\Psi}_2 &= -\mathbf{I}.\end{aligned}$$

After direct multiplication of these equations by $-\mathbf{I}$, $-\mathbf{\Phi}$, $-\mathbf{\Phi}^2$, $-\mathbf{\Phi}^3$, respectively, and adding, we find the equation:

$$\mathbf{\Phi}^3 - \Phi_1\mathbf{\Phi}^2 + \Phi_2\mathbf{\Phi} - \Phi_3\mathbf{I} = 0. \tag{9}$$

This equation is called the Hamilton-Cayley Equation, Hamilton having shown that a quaternion satisfies a corresponding equation, while Cayley showed that any square matrix of the n'th order satisfies a corresponding characteristic equation of the same order.

With the aid of the Hamilton-Cayley equation and the corresponding scalar equation (4) in λ, a systematic classification of the various forms which a dyadic can assume is possible.

§71

On the Classification of Dyadics

It was shown in Art. 70 that any dyadic $\mathbf{\Phi}$ satisfies identically the Hamilton-Cayley equation:

$$\mathbf{\Phi}^3 - \Phi_1\mathbf{\Phi}^2 + \Phi_2\mathbf{\Phi} - \Phi_3\mathbf{I} = 0,$$

where the scalar invariants Φ_1, Φ_2, Φ_3 are the same as those in the corresponding scalar equation:

$$\lambda^3 - \Phi_1\lambda^2 + \Phi_2\lambda - \Phi_3 = 0.$$

It may happen, however, that $\mathbf{\Phi}$ is of such a nature as to satisfy identically an equation of the second or first degree. In any case, the equation of lowest degree satisfied by a dyadic $\mathbf{\Phi}$, together with the roots of the corresponding scalar equation for λ, can be used as a basis for its classification. These equations are called, respectively, the Characteristic Equation for $\mathbf{\Phi}$, and the Characteristic Equation for λ.

Let l, m, n denote the three roots of the general equation for λ; then the Hamilton-Cayley equation can be written:

$$(\mathbf{\Phi} - l\mathbf{I})\cdot(\mathbf{\Phi} - m\mathbf{I})\cdot(\mathbf{\Phi} - n\mathbf{I}) = 0,$$

whatever the values of l, m, n.

There are three generic classes to one of which any dyadic Φ must belong:

(A) Φ satisfies identically an equation of the third degree and none of lower degree.

(B) Φ satisfies identically an equation of the second degree and none of lower degree.

(C) Φ satisfies identically an equation of the first degree.

Under A there are four sub-classes into one of which any dyadic Φ of class A must fall, and which are characterized by the forms assumed by the Hamilton-Cayley equation, under the possible assumptions as to the nature of the roots l, m, n, viz:

(A-I) $$(\Phi - l\mathrm{I})\cdot(\Phi - m\mathrm{I})\cdot(\Phi - n\mathrm{I}) = 0,$$

with l, m, n all real and unequal;

(A-II) $$(\Phi - l\mathrm{I})\cdot[\Phi^2 - 2\alpha\Phi + (\alpha^2 + \beta^2)\,\mathrm{I}] = 0,$$

with l real, $m = \alpha + \beta i$, $n = \alpha - \beta i$, and $\beta \neq 0$; $i = \sqrt{-1}$;

(A-III) $$(\Phi - l\mathrm{I})\cdot(\Phi - m\mathrm{I})^2 = 0,$$

with l, m, n all real, and $m = n \neq l$;

(A-IV) $$(\Phi - l\mathrm{I})^3 = 0,$$

with l, m, n all real and equal.

Referring now to class B, let $F(\Phi) = 0$ represent the Hamilton-Cayley equation, and $G(\Phi) = 0$ the characteristic equation for Φ. Then we can write:

$$F(\Phi) = G(\Phi)\cdot Q(\Phi) + R(\Phi) = 0;$$

and also the corresponding scalar equation:

$$F(\lambda) = G(\lambda)\,Q(\lambda) + R(\lambda) = 0;$$

where Q and R represent linear functions. Since $F(\Phi)$ and $G(\Phi)$ must vanish identically, and since in accordance with hypothesis an equation in Φ of degree lower than the second does not exist, it is necessary that $R(\Phi)$, and therefore $R(\lambda)$, shall vanish identically. It follows that, for the generic class B:

$$F(\Phi) = G(\Phi)\cdot Q(\Phi) = 0,$$
$$F(\lambda) = G(\lambda)\,Q(\lambda) \quad = 0.$$

The roots of $G(\lambda) = 0$ must therefore be included among those of $F(\lambda) = 0$. Hence, denoting the common roots of these two equations by l and m, we have:

$$G(\lambda) = (\lambda - l)\,(\lambda - m).$$

Under B there are two sub-classes, characterized by special forms of the characteristic equation, $G(\Phi) = 0$, and of the roots of $G(\lambda) = 0$, which we are now able to specify as follows:

$$\text{(B-I)} \qquad (\Phi - lI) \cdot (\Phi - mI) = 0,$$

with l and m both real or complex and $l \neq m$.

$$\text{(B-II)} \qquad (\Phi - lI)^2 = 0,$$

with l real and $l = m$.

Referring now to class C, and following the method used in discussing class B, we find:

$$F(\Phi) = G(\Phi) \cdot Q(\Phi) = 0,$$
$$F(\lambda) = G(\lambda)\, Q(\lambda) \quad = 0,$$

where $G(\Phi) = 0$ and $G(\lambda) = 0$, respectively, now denote the equations of the first degree satisfied by Φ and λ; and the root of $G(\lambda) = 0$, denoted by l, must be one of the roots of $F(\lambda) = 0$. Hence:

$$G(\lambda) = (\lambda - l).$$

Under C there is, consequently, but one class of dyadics, viz: such as satisfy identically the characteristic equation:

$$\text{(C-I)} \qquad \Phi - lI = 0.$$

There are, then, seven essentially different forms of dyadics which satisfy respectively the seven characteristic equations A-I, II, III, IV, B-I, II, and C-I. These, with the names given them by Gibbs-Wilson, can be expressed as follows:

(A_I)	$l\mathrm{a}_1\mathrm{a}^1 + m\mathrm{a}_2\mathrm{a}^2 + n\mathrm{a}_3\mathrm{a}^3,$	Tonic;
(A_{II})	$l\mathrm{a}_1\mathrm{a}^1 + \alpha(\mathrm{a}_2\mathrm{a}^2 + \mathrm{a}_3\mathrm{a}^3) + \beta(\mathrm{a}_3\mathrm{a}^2 - \mathrm{a}_2\mathrm{a}^3),$	Cyclotonic;
(A_{III})	$l\mathrm{a}_1\mathrm{a}^1 + m(\mathrm{a}_2\mathrm{a}^2 + \mathrm{a}_3\mathrm{a}^3) + \mathrm{a}_3\mathrm{a}^2,$	Simple Shearer;
(A_{IV})	$lI + \mathrm{a}_1\mathrm{a}^2 + \mathrm{a}_2\mathrm{a}^3,$	Complex Shearer;
(B_I)	$l\mathrm{a}_1\mathrm{a}^1 + m(\mathrm{a}_2\mathrm{a}^2 + \mathrm{a}_3\mathrm{a}^3) + \mathrm{a}_3\mathrm{a}^2,$	Special Tonic;
(B_{II})	$lI + \mathrm{a}_3\mathrm{a}^2,$	Special Simple Shearer;
(C_I)	$lI,$	Special Tonic;

where $\mathrm{a}_1, \mathrm{a}_2, \mathrm{a}_3$ and $\mathrm{a}^1, \mathrm{a}^2, \mathrm{a}^3$ denote reciprocal systems of vectors, l, m, n roots of the characteristic equation for λ, and α and β are given by the equations $m = \alpha + \beta i$ and $n = \alpha - \beta i$, when m and n are complex. Each of these dyadics when substituted for Φ in the corresponding equation can be seen to satisfy it identically.

Conversely, it can be shown that, if any dyadic $\mathbf{\Phi}$ is known to satisfy any one of the seven characteristic equations, then it must be capable of reduction to the corresponding form in the above list. The proof of this statement is somewhat long and tedious. It can be found in Chapter VI of the treatise on Vector Analysis by Gibbs-Wilson.

§72
Applications of Dyadics

(a) Equations of quadric surfaces. If $\mathbf{r}$ denote the position vector of a point P with respect to an origin O, and $\mathbf{\Phi}$ any dyadic, then, as will be shown, the equations:

(1a) $$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 1,$$

(2a) $$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 0,$$

are respectively the dyadic equations of a general quadric surface, real or imaginary, with center at O, and of a cone, real or imaginary, with vertex at O.

In these equations the dyadic $\mathbf{\Phi}$ may be considered as symmetric; for, in any case it can, as seen in Art. 66, be expressed as the sum of a symmetric and an anti-symmetric part, and, if $\mathbf{\Psi}$ denote the anti-symmetric part, the contributions of this part to the left-hand members will be $\mathbf{r} \cdot \mathbf{\Psi} \cdot \mathbf{r}$, but, since $\mathbf{\Psi}$ is anti-symmetric, $\mathbf{\Psi} \cdot \mathbf{r}$ must be a vector perpendicular to $\mathbf{r}$, and hence $\mathbf{r} \cdot \mathbf{\Psi} \cdot \mathbf{r} = 0$.

We suppose, then, that $\mathbf{\Phi}$ is symmetric, and also suppose it expressed in accordance with the procedure of Art. 68 in the normal form:

(3a) $$\mathbf{\Phi} = \pm \frac{\mathbf{ii}}{a^2} \pm \frac{\mathbf{jj}}{b^2} \pm \frac{\mathbf{kk}}{c^2},$$

where a^2, b^2, c^2 are positive constants. Let $\mathbf{r}$ be resolved into components as follows:

(4a) $$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k},$$

where x, y, z are rectangular co-ordinates of the point P. We then find:

(5a) $$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = \pm \frac{x^2}{a^2} \pm \frac{y^2}{b^2} \pm \frac{z^2}{c^2}.$$

Hence, the equation

$$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 1$$

represents:

a real ellipsoid, if all signs in equation (3a) are positive;
an hyperboloid of one sheet, if one sign is negative;
an hyperboloid of two sheets, if two signs are negative;
an imaginary quadric, if all signs are negative;

and the equation

$$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 0,$$

represents:

a real cone, if the signs in equation (3a) are not all the same;
an imaginary cone, if the signs are all the same.

For a discussion of the properties of quadric surfaces with the aid of dyadics, the reader is referred to Vector-Analysis, Gibbs-Wilson, p. 372.

(b) Representation of an affine transformation by a dyadic. It was shown in Art. 56 that an affine transformation can be represented by a linear vector function. Such a transformation must, therefore, also be capable of representation by a dyadic.

An affine transformation with origin fixed is expressed by the equations:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ y_3 &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{aligned} \tag{1b}$$

where the a-coefficients are scalar parameters and x_1, x_2, x_3 and y_1, y_2, y_3, respectively, are the affine co-ordinates on an $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$-base-system of a point P and of the point Q into which P goes in the transformation.

The dyadic

$$\begin{aligned} \mathbf{\Phi} = \quad & a_{11}\mathfrak{a}_1\mathfrak{a}^1 + a_{12}\mathfrak{a}_1\mathfrak{a}^2 + a_{13}\mathfrak{a}_1\mathfrak{a}^3 \\ & + a_{21}\mathfrak{a}_2\mathfrak{a}^1 + a_{22}\mathfrak{a}_2\mathfrak{a}^2 + a_{23}\mathfrak{a}_2\mathfrak{a}^3 \\ & + a_{31}\mathfrak{a}_3\mathfrak{a}^1 + a_{32}\mathfrak{a}_3\mathfrak{a}^2 + a_{33}\mathfrak{a}_3\mathfrak{a}^3 \end{aligned} \tag{2b}$$

will represent the transformation. For, acting as a prefactor upon the position-vector of P, viz:

$$\mathbf{r} = x_1\mathfrak{a}_1 + x_2\mathfrak{a}_2 + x_3\mathfrak{a}_3,$$

it produces the position-vector of Q, viz:

$$\mathbf{q} = y_1\mathfrak{a}_1 + y_2\mathfrak{a}_2 + y_3\mathfrak{a}_3,$$

where y_1, y_2, y_3 are expressed in terms of x_1, x_2, x_3 by equations (1b).

If $\mathbf{\Psi}$ be a dyadic representing a second affine transformation whereby the point Q is further transformed into the point Q' having the position vector q', then:

$$\mathrm{q}' = \mathbf{\Psi} \cdot \mathrm{q} = \mathbf{\Psi} \cdot (\mathbf{\Phi} \cdot \mathbf{r}),$$

and hence, since the latter product is associative:

$$\mathrm{q}' = \mathbf{\Psi} \cdot (\mathbf{\Phi} \cdot \mathbf{r}) = (\mathbf{\Psi} \cdot \mathbf{\Phi}) \cdot \mathbf{r}. \tag{3b}$$

From the last equation the important conclusion can be drawn that the transformation effected by the dyadic $\mathbf{\Phi}$ followed by the transformation effected by the dyadic $\mathbf{\Psi}$ is equivalent to the single transformation effected by the dyadic $\mathbf{\Psi} \cdot \mathbf{\Phi}$. We shall have occasion to refer again to this result in Art. 83, which deals with the affine transformation group.

(c) **Rotation and perversion expressed by dyadics.**

Consider an affine transformation which consists of a rotation of the points of space through the same angle θ about an axis passing through a fixed point O and determined by a unit vector $\mathbf{a}$. It is required to find a dyadic $\mathbf{\Phi}$ which acting as a prefactor will express the rotation.

Let $\mathbf{r}$ denote a unit vector perpendicular to $\mathbf{a}$; then $\mathbf{a} \times \mathbf{r}$ will be a unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{r}$. Since $\mathbf{a}$ must remain unaltered by the rotation, and since $\mathbf{r}$ must be rotated through the angle θ in its plane perpendicular to $\mathbf{a}$, the required dyadic $\mathbf{\Phi}$ must be such that:

$$\begin{aligned} \mathbf{\Phi} \cdot \mathbf{a} &= \mathbf{a}; \\ \mathbf{\Phi} \cdot \mathbf{r} &= \cos \theta \mathbf{r} + \sin \theta\, \mathbf{a} \times \mathbf{r}; \end{aligned}$$

and these conditions are necessary and sufficient that $\mathbf{\Phi}$ shall represent the rotation. The conditions can be fulfilled by taking:

$$\mathbf{\Phi} = \mathbf{a}\mathbf{a} + \cos \theta(\mathrm{I} - \mathbf{a}\mathbf{a}) + \sin \theta \mathrm{I} \times \mathbf{a}, \tag{1c}$$

where I represents an idemfactor, as is evident upon trial.

The dyadic $\mathbf{\Phi}$ can easily be expressed in a nonian form in which all the given data appear explicitly. Let $\mathbf{a}$ be expressed in terms of its components on an arbitrary **i**, **j**, **k**-base-system as follows:

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k},$$

where the measure-numbers a_1, a_2, a_3 of the components are to be considered as given; the origin is assumed at the fixed point O. With respect to this base-system we shall then have:

$$
\begin{aligned}
\mathfrak{a}\mathfrak{a} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\\
&= \quad a_1a_1\mathbf{ii} + a_1a_2\mathbf{ij} + a_1a_3\mathbf{ik}\\
&\quad + a_2a_1\mathbf{ji} + a_2a_2\mathbf{jj} + a_2a_3\mathbf{jk}\\
&\quad + a_3a_1\mathbf{ki} + a_3a_2\mathbf{kj} + a_3a_3\mathbf{kk};\\
\mathbf{I} &= \quad \mathbf{ii} + \mathbf{jj} + \mathbf{kk};\\
\mathbf{I}\times\mathfrak{a} &= \quad (\mathbf{ii} + \mathbf{jj} + \mathbf{kk})\times(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\\
&= \quad 0\mathbf{ii} - a_3\mathbf{ij} + a_2\mathbf{ik}\\
&\quad + a_3\mathbf{ji} + 0\mathbf{jj} - a_1\mathbf{jk}\\
&\quad - a_2\mathbf{ki} + a_1\mathbf{kj} + 0\mathbf{kk}.
\end{aligned}
$$

Introducing these expressions into equation (1c), we find:

$$
\begin{aligned}
\mathbf{\Phi} = &\quad \{a_1^2(1-\cos\theta) + \cos\theta\}\mathbf{ii}\\
&+\{a_1a_2(1-\cos\theta) - a_3\sin\theta\}\mathbf{ij}\\
&+\{a_1a_3(1-\cos\theta) + a_2\sin\theta\}\mathbf{ik}\\
&+\{a_2a_1(1-\cos\theta) + a_3\sin\theta\}\mathbf{ji}\\
&+\{\,{}_2^2(1-\cos\theta) + \cos\theta\}\mathbf{jj}\\
&+\{a_2a_3(1-\cos\theta) - a_1\sin\theta\}\mathbf{jk}\\
&+\{a_3a_1(1-\cos\theta) - a_2\sin\theta\}\mathbf{ki}\\
&+\{a_3a_2(1-\cos\theta) + a_1\sin\theta\}\mathbf{kj}\\
&+\{a_3^2(1-\cos\theta) + \cos\theta\}\mathbf{kk},
\end{aligned}
\tag{2c}
$$

which is the required nonian form for $\mathbf{\Phi}$. In this expression a_1, a_2, a_3, and θ are given parameters, but the first three of these are direction cosines of the axis of rotation and, therefore, are subject to the relation $a_1^2 + a_2^2 + a_3^2 = 1$; to specify a rotation three parameters only are necessary, two to specify the direction of the axis, and one to specify the angle of rotation.

The dyadic $\mathbf{\Phi}$ can also be expressed in the following trinomial form:

$$\mathbf{\Phi} = \mathbf{i}'\mathbf{i} + \mathbf{j}'\mathbf{j} + \mathbf{k}'\mathbf{k}, \tag{3c}$$

where:

$$\mathbf{i}' = \mathbf{\Phi}\cdot\mathbf{i}, \qquad \mathbf{j}' = \mathbf{\Phi}\cdot\mathbf{j}, \qquad \mathbf{k}' = \mathbf{\Phi}\cdot\mathbf{k}.$$

The vectors $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ are obviously the unit vectors of the right-handed congruent system into which the $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-system is transformed by the rotation.

By comparison of equations (2c) and (3c) we find:

$$
\begin{aligned}
\mathbf{i}' = &\quad \{a_1^2(1-\cos\theta) + \cos\theta\}\mathbf{i}\\
&+\{a_2a_1(1-\cos\theta) + a_3\sin\theta\}\mathbf{j}\\
&+\{a_3a_1(1-\cos\theta) - a_2\sin\theta\}\mathbf{k},
\end{aligned}
\tag{4c}
$$

and there are corresponding expressions for $\mathbf{j}'$ and $\mathbf{k}'$ obtained by cyclical permutation of the subscripts 1, 2, 3 and of the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.

Any dyadic reducible to the form (3c) is called a Versor. It may be observed that the conjugate of the dyadic $\mathbf{\Phi}$ acting as a prefactor will turn the unit vectors $\mathbf{i}', \mathbf{j}', \mathbf{k}'$ into the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$. It follows that a versor and its conjugate must be reciprocal dyadics.

By equations (6), (7), and (8), Art. 69, the following quantities must be versor invariants:

$$\begin{aligned} \Phi_1 &= \mathbf{i}' \cdot \mathbf{i} + \mathbf{j}' \cdot \mathbf{j} + \mathbf{k}' \cdot \mathbf{k}, \\ \Phi_2 &= \mathbf{j}' \times \mathbf{k}' \cdot \mathbf{i} + \mathbf{k}' \times \mathbf{i}' \cdot \mathbf{j} + \mathbf{i}' \times \mathbf{j}' \cdot \mathbf{k}, \\ \Phi_3 &= \mathbf{i}' \cdot \mathbf{j}' \times \mathbf{k}'. \end{aligned} \tag{5c}$$

It is seen at once that the first two of these quantities are equal and that the third is equal to unity. Since they are invariants, they must retain the same values however the arbitrary $\mathbf{i}, \mathbf{j}, \mathbf{k}$-base-system be chosen. If we take the unit vector $\mathbf{i}$ to coincide with the unit vector $\mathbf{a}$, then:

$$\mathbf{i}' \cdot \mathbf{i} = 1, \quad \mathbf{j}' \cdot \mathbf{j} = \mathbf{k}' \cdot \mathbf{k} = \cos\theta.$$

Hence:

$$\Phi_1 = \Phi_2 = 1 + 2\cos\theta; \qquad \Phi_3 = 1. \tag{6c}$$

By equation (8′), Art. 69, the invariant Φ_3 is equal to the determinant of the coefficients of the dyads in the nonian form for $\mathbf{\Phi}$ given by equation (2c), and since $\Phi_3 = 1$, this determinant must have the value unity for any arbitrary choice of the $\mathbf{i}, \mathbf{j}, \mathbf{k}$-base-system.

If $\mathbf{\Psi}$ denote the negative of the versor $\mathbf{\Phi}$, then:

$$\mathbf{\Psi} = -(\mathbf{i}'\mathbf{i} + \mathbf{j}'\mathbf{j} + \mathbf{k}'\mathbf{k}); \tag{7c}$$

the dyadic $\mathbf{\Psi}$ is called a Perversor. Acting upon the position-vectors of any figure it rotates the figure and then changes it into another symmetrical to itself with the origin as a point of symmetry. This sort of transformation is called a Perversion.

If in the general expression for a versor given by equation (1c) the angle θ be taken as $\pi/2$, we get:

$$\mathbf{\Phi} = \mathbf{aa} + \mathbf{I} \times \mathbf{a}. \tag{8c}$$

This dyadic is called a Quadrantal Versor. Acting as a prefactor upon any vector perpendicular to an axis determined by $\mathbf{a}$, it turns the vector through one right angle.

Again, if in equation (1c) the angle θ be taken as π, we get:

$$\mathbf{\Phi} = 2\mathbf{aa} - \mathbf{I}. \tag{9c}$$

This dyadic is called a Bi-Quadrantal Versor. Acting as a prefactor upon any vector perpendicular to an axis determined by $\mathbf{a}$ it turns the vector through two right angles.

The nonian forms for a quadrantal and for a bi-quadrantal versor can be obtained from equation (2c) by the substitution of $\pi/2$ and π, respectively, for the angle θ.

(d) **The inertia dyadic for a system of mass particles; principal moments and axes of inertia.** Let $\mathbf{r}$ denote the position-vector with respect to an arbitrary origin O of a typical particle the magnitude of whose mass is denoted by m, and let $\boldsymbol{\rho}$ denote the position-vector of any point P with respect to O, and p the magnitude of the distance of the typical particle from a line through O in the direction of $\boldsymbol{\rho}$. Then, if K denote the magnitude of the moment of inertia of the system, with respect to this line, and if Σ indicate summation over all the particles:

$$K = \Sigma m p^2,$$

or, making use of the Pythagorean theorem:

$$K = \Sigma m\left(\mathbf{r} \cdot \mathbf{r} - \frac{\mathbf{r} \cdot \boldsymbol{\rho}\, \mathbf{r} \cdot \boldsymbol{\rho}}{\boldsymbol{\rho} \cdot \boldsymbol{\rho}}\right),$$

an equation which, after easy transformation, can be written in the form:

$$K = \frac{1}{\rho^2}\boldsymbol{\rho} \cdot [\Sigma m(\mathbf{r} \cdot \mathbf{r}\mathbf{I} - \mathbf{r}\mathbf{r})] \cdot \boldsymbol{\rho},$$

or, more simply, in the form:

(1d)
$$K = \frac{\boldsymbol{\rho} \cdot \boldsymbol{\Phi} \cdot \boldsymbol{\rho}}{\rho^2},$$

where:

(2d)
$$\boldsymbol{\Phi} = \Sigma m(\mathbf{r} \cdot \mathbf{r}\mathbf{I} - \mathbf{r}\mathbf{r}).$$

This dyadic is called the Inertia Dyadic of the system of mass particles.

Suppose, now, that $\boldsymbol{\rho}$ is the position-vector of any point P on the quadric surface whose equation is:

(3d)
$$\boldsymbol{\rho} \cdot \boldsymbol{\Phi} \cdot \boldsymbol{\rho} = 1.$$

Since K must be finite for all possible values of $\boldsymbol{\rho}$, it follows from equations (1d) and (3d) that $\boldsymbol{\rho}$ must always be finite, and hence that the quadric surface must be an ellipsoid. It is called the Poinsot Ellipsoid of Inertia.

By equations (1d) and (3d) this ellipsoid has the important property expressed by the equation:

(4d)
$$K = \frac{1}{\rho^2},$$

which implies that the moment of inertia of the system of mass particles with respect to a line through O collinear with any radius vector of the ellipsoid varies inversely with the square of the radius vector.

The inertia dyadic can be expressed in nonian form by introducing an $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$-base-system with origin at O, and writing:

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \qquad \mathbf{I} = \mathbf{ii} + \mathbf{jj} + \mathbf{kk}.$$

From equation (2d) we then get:

$$\begin{array}{rl} \text{(5d)} \quad \mathbf{\Phi} = & [\Sigma m(y^2 + z^2)]\,\mathbf{ii} - [\Sigma mxy]\,\mathbf{ij} - [\Sigma mxz]\,\mathbf{ik} \\ & -[\Sigma myx]\,\mathbf{ji} + [\Sigma m(z^2 + x^2)]\,\mathbf{jj} - [\Sigma myz]\,\mathbf{jk} \\ & -[\Sigma mzx]\,\mathbf{ki} - [\Sigma mzy]\,\mathbf{kj} + [\Sigma m(x^2 + y^2)]\,\mathbf{kk}. \end{array}$$

This form for the inertia dyadic shows it to be symmetric. The coefficients in the principal diagonal of this form are the magnitudes of the moments of inertia of the system of particles with respect to the co-ordinate axes, and the negatives of the other coefficients are those of products of inertia with respect to co-ordinate planes.

Since the inertia dyadic is symmetric, it can be reduced to the normal form:

$$\text{(6d)} \qquad \mathbf{\Phi} = K_1\mathbf{i}_0\mathbf{i}_0 + K_2\mathbf{j}_0\mathbf{j}_0 + K_3\mathbf{k}_0\mathbf{k}_0,$$

where $\mathbf{i}_0$, $\mathbf{j}_0$, $\mathbf{k}_0$ represent a special choice of $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, and where K_1, K_2, K_3, are the magnitudes of the moments of inertia with respect to axes determined by $\mathbf{i}_0$, $\mathbf{j}_0$, $\mathbf{k}_0$. These axes are called Principal Axes of Inertia, and the corresponding moments of inertia the Principal Moments of Inertia with respect to these axes.

(e) The strain dyadic for small strain in an elastic continuum. The subject of strain in an elastic continuum is one which is particularly well adapted to show the advantages of dyadic methods.

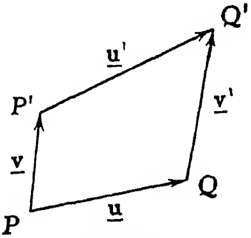


Fig. 35.

Referring to Fig. 35, let P denote any point of an elastic continuum, and Q any neighboring point at a small distance from P. Suppose the continuum to undergo a small strain so that P moves to P' and Q to Q', and let $\mathbf{u}$, $\mathbf{u}'$, $\mathbf{v}$, and $\mathbf{v}'$ denote vectors represented, respectively, by the line-vectors PQ, $P'Q'$, PP', and QQ'.

The vector $\mathbf{s}$, defined by the equation:

$$\text{(1e)} \qquad \mathbf{s} = \mathbf{u}' - \mathbf{u},$$

is called the Shift of the vector **u** due to the strain. It is proposed to show that **s** is a linear vector function of **u**, or, which amounts to the same thing, that:

$$\mathbf{s} = \boldsymbol{\Phi} \cdot \mathbf{u},$$

where $\boldsymbol{\Phi}$ represents a dyadic, called the Strain Dyadic, which, when specified, determines the nature of the strain.

Referring again to Fig. 35, it is seen that:

$$\mathbf{u}' = -\mathbf{v} + \mathbf{u} + \mathbf{v}'.$$

Hence:

$$\mathbf{s} = \mathbf{u}' - \mathbf{u} = \mathbf{v}' - \mathbf{v}. \tag{2e}$$

We suppose the strain to be continuous, so that the vector **v** can be regarded as a continuous function of position, and we refer the **u** and **v**-vectors to a fixed **i**, **j**, **k**-system, so that:

$$\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k},$$
$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k},$$

where the u and v-coefficients are the measure-numbers of the components of **u** and **v**.

Then, to terms of the second order in the small quantities u_1, u_2, u_3, we have, with the aid of Taylor's theorem:

$$\begin{aligned} & \mathbf{v} + \left(\frac{\partial v_1}{\partial x} u_1 + \frac{\partial v_1}{\partial y} u_2 + \frac{\partial v_1}{\partial z} u_3\right) \mathbf{i} \\ & + \left(\frac{\partial v_2}{\partial x} u_1 + \frac{\partial v_2}{\partial y} u_2 + \frac{\partial v_2}{\partial z} u_3\right) \mathbf{j} \\ & + \left(\frac{\partial v_3}{\partial x} u_1 + \frac{\partial v_3}{\partial y} u_2 + \frac{\partial v_3}{\partial z} u_3\right) \mathbf{k}, \end{aligned} \tag{3e}$$

where the nine partial derivatives are supposed evaluated at the point P.

We now define a dyadic $\boldsymbol{\Phi}$ with these nine partial derivatives as coefficients:

$$\begin{aligned} & \frac{\partial v_1}{\partial x} \mathbf{ii} + \frac{\partial v_1}{\partial y} \mathbf{ij} + \frac{\partial v_1}{\partial z} \mathbf{ik} \\ & + \frac{\partial v_2}{\partial x} \mathbf{ji} + \frac{\partial v_2}{\partial y} \mathbf{jj} + \frac{\partial v_2}{\partial z} \mathbf{jk} \\ & + \frac{\partial v_3}{\partial x} \mathbf{ki} + \frac{\partial v_3}{\partial y} \mathbf{kj} + \frac{\partial v_3}{\partial z} \mathbf{kk}. \end{aligned} \tag{4e}$$

The equation for $\mathbf{v}'$ can now be written:

$$\mathbf{v}' = \mathbf{v} + \boldsymbol{\Phi} \cdot \mathbf{u}. \tag{5e}$$

This equation with (2e) gives:

$$\mathbf{s} = \boldsymbol{\Phi} \cdot \mathbf{u}. \tag{6e}$$

This is a dyadic equation specifying the strain in the immediate neighborhood of the point P.

If the coefficients of $\mathbf{\Phi}$ are constants, that is, independent of the position of the point P, the continuum is said to have undergone a Homogeneous Strain, in which, as can easily be shown, straight lines go into straight lines, and parallel lines go into parallel lines. In any case, a continuous strain in an elastic continuum may be considered as homogeneous in the immediate neighborhood of any specified point.

We now define nine strain parameters by the equations:

$$\text{(7e)}\qquad \begin{aligned} a_1 &= \frac{\partial v_1}{\partial x}, & g_1 &= \frac{1}{2}\left(\frac{\partial v_2}{\partial z} + \frac{\partial v_3}{\partial y}\right), & \omega_1 &= \frac{1}{2}\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right), \\ a_2 &= \frac{\partial v_2}{\partial y}, & g_2 &= \frac{1}{2}\left(\frac{\partial v_3}{\partial x} + \frac{\partial v_1}{\partial z}\right), & \omega_2 &= \frac{1}{2}\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right), \\ a_3 &= \frac{\partial v_3}{\partial z}, & g_3 &= \frac{1}{2}\left(\frac{\partial v_1}{\partial y} + \frac{\partial v_2}{\partial x}\right), & \omega_3 &= \frac{1}{2}\left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right); \end{aligned}$$

the a-parameters are called Elongation Coefficients, the g-parameters Shear Coefficients, and the ω-parameters Rotation Coefficients.

The strain dyadic can now be written as the sum of two dyadics:

$$\text{(8e)}\qquad \mathbf{\Phi} = \mathbf{\Psi} + \mathbf{\Omega},$$

where:

$$\text{(9e)}\qquad \begin{aligned} \mathbf{\Psi} = {}& a_1\mathbf{ii} + g_3\mathbf{ij} + g_2\mathbf{ik} \\ & + g_3\mathbf{ji} + a_2\mathbf{jj} + g_1\mathbf{jk} \\ & + g_2\mathbf{ki} + g_1\mathbf{kj} + a_3\mathbf{kk}, \end{aligned}$$

$$\text{(10e)}\qquad \begin{aligned} \mathbf{\Omega} = {}& 0 - \omega_3\mathbf{ij} + \omega_2\mathbf{ik} \\ & + \omega_3\mathbf{ji} + 0 - \omega_1\mathbf{jk} \\ & - \omega_2\mathbf{ki} + \omega_1\mathbf{kj} + 0. \end{aligned}$$

If: $\mathbf{\Psi} = 0$, $\mathbf{\Omega} \neq 0$, the strain can be shown to consist of a rotation only; if: $\mathbf{\Psi} \neq 0$, $\mathbf{\Omega} = 0$, the strain is called a Pure Strain.

To proceed further with the development of the theory of strain would lead us beyond the scope of the present book. But sufficient has been said, perhaps, to indicate the usefulness of dyadic methods in dealing with this branch of mathematical physics.

It may be remarked, however, before leaving this subject, that, in virtue of the fundamental strain equation:

$$\mathbf{s} = \mathbf{\Phi} \cdot \mathbf{u},$$

the essentially different types of strain in an elastic continuum can be specified by the essentially different types of dyadics which have been classified and discussed in Art. 71.

The stress which is associated with small strain can also be described by the method of dyadics as will be seen later (Art. 109), when the Stress Dyadic is introduced as an example of a Tensor.

(f) Fields of electric force and polarization in a crystal.

As a final example of the utility of dyadics in dealing with physical problems we shall consider the relationship of the fields of electric force and electric polarization in a crystal.

Electrical theory shows that the polarization can be expressed in terms of the force with the aid of nine parameters, only six of which, however, are independent.

Let **E** and **P** respectively, denote the force and polarization vectors, and let these vectors be resolved into their components on an arbitrary **i**, **j**, **k**-base-system as follows:

$$\mathbf{E} = E_1\mathbf{i} + E_2\mathbf{j} + E_3\mathbf{k}; \tag{1f}$$
$$\mathbf{P} = P_1\mathbf{i} + P_2\mathbf{j} + P_3\mathbf{k}; \tag{2f}$$

where the E and P coefficients, respectively are the measure-numbers of the components.

By considering the polarized molecules of the crystal as electric doublets whose axes, owing to the crystalline structure, do not coincide in direction with the electric field which excites them, the theory indicates that the measure-numbers of the force and polarization components are related as follows:

$$\begin{aligned} P_1 &= a_{11}E_1 + a_{12}E_2 + a_{13}E_3, \\ P_2 &= a_{21}E_1 + a_{22}E_2 + a_{23}E_3, \\ P_3 &= a_{31}E_1 + a_{32}E_2 + a_{33}E_3, \end{aligned} \tag{3f}$$

where the a-coefficients are parameters whose numerical values depend upon the nature of the crystal and upon the system of units which is used.

We now form a dyadic:

$$\begin{aligned} \mathbf{\Phi} = \; & a_{11}\mathbf{ii} + a_{12}\mathbf{ij} + a_{13}\mathbf{ik} \\ & + a_{21}\mathbf{ji} + a_{22}\mathbf{jj} + a_{23}\mathbf{jk} \\ & + a_{31}\mathbf{ki} + a_{32}\mathbf{kj} + a_{33}\mathbf{kk}. \end{aligned} \tag{4f}$$

If we operate with this dyadic as a prefactor upon the electric force vector as expressed by equation (1f), we find:

$$\mathbf{P} = \mathbf{\Phi} \cdot \mathbf{E}. \tag{5f}$$

The electric polarization vector is, therefore, a linear vector function of the electric force vector, which can be obtained from it by operating with the dyadic $\mathbf{\Phi}$ as a prefactor.

EXERCISES ON CHAPTER VI

1. Prove the following statements:
   (a) If $\mathbf{\Phi} \cdot \mathbf{r} = 0$, for three non-coplanar values of $\mathbf{r}$, then $\mathbf{\Phi} = 0$.
   (b) If $\mathbf{\Phi} \cdot \mathbf{r} = 0$, for two non-collinear values of $\mathbf{r}$, then the vectors $\mathbf{\Phi} \cdot \mathbf{r}$ which do not vanish must be parallel.
   (c) If $\mathbf{\Phi} \cdot \mathbf{r} = 0$, for a given value of $\mathbf{r}$, then the vectors $\mathbf{\Phi} \cdot \mathbf{r}$ which do not vanish must be coplanar.

2. Show that $\mathbf{\Phi} \cdot \mathbf{\Phi}_c$ is a symmetric dyadic.

3. Show that:

$$(\mathbf{I} \times \mathbf{a}) \cdot \mathbf{\Phi} = \mathbf{a} \times \mathbf{\Phi},$$
$$(\mathbf{a} \times \mathbf{I}) \cdot \mathbf{\Phi} = \mathbf{a} \times \mathbf{\Phi}.$$

4. If $\mathbf{\Omega}$ is a complete dyadic and if $\mathbf{\Phi} \cdot \mathbf{\Omega} = \mathbf{\Psi} \cdot \mathbf{\Omega}$, show that $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are equal.

5. Prove that the necessary and sufficient condition that an anti-symmetric dyadic $\mathbf{\Phi}$ shall vanish is that the vector of $\mathbf{\Phi}$ shall vanish

6. Show that:

$$(\mathbf{\Phi} \times \mathbf{a})_c = -\mathbf{a} \times \mathbf{\Phi}_c.$$

7. If $\mathbf{\Phi}$ is small, show that $\mathbf{I} + \mathbf{\Phi}$ and $\mathbf{I} - \mathbf{\Phi}$ can be considered as reciprocal dyadics.

8. If $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are vectors representing conjugate radii of an ellipsoid whose equation is:

$$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 1,$$

show that the equation of a plane through the terminal points of the three conjugate radii is:

$$\mathbf{r} \cdot \mathbf{\Phi} \cdot (\mathbf{a} + \mathbf{b} + \mathbf{c}) = 1.$$

9. If $\mathbf{r}' = \mathbf{\Phi} \cdot \mathbf{r}$ and if $\mathbf{r}$ is the position-vector with respect to its center of a generic point of an ellipsoid whose equation is:

$$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 1,$$

show that the locus of the terminal point of a line-vector drawn from the center and representing $\mathbf{r}'$ is an ellipsoid (the reciprocal ellipsoid).

10. Prove that the necessary and sufficient condition that two quadrics whose equations are:

$$\mathbf{r} \cdot \mathbf{\Phi} \cdot \mathbf{r} = 1, \qquad \mathbf{r} \cdot \mathbf{\Psi} \cdot \mathbf{r} = 1,$$

shall be confocal is that $\mathbf{\Phi}$ and $\mathbf{\Psi}$ differ only by a scalar multiple of the idemfactor.

11. If the reciprocal of a dyadic and its conjugate are equal, show that it must be a versor or perversor.

12. Show that any versor can be expressed as the direct product of two quadrantal versors whose axes lie in the plane perpendicular to the axis of the versor.

CHAPTER VII

CO-ORDINATE SYSTEMS

§73

Introduction of General Co-ordinates
Unitary and Reciprocal Unitary Vectors

In many problems relating to vector fields it is often advantageous to use co-ordinate systems of greater generality than those hitherto employed, and of which the latter are but special cases. It is proposed, therefore, in the present chapter to discuss in some detail the properties of various types of co-ordinate systems.

The notation which will be used in this chapter is purposely made consistent with the covariant and contravariant notation, which is explained in the next chapter.

Let u^1, u^2, u^3 denote three independent, continuous, and single-valued scalar point functions between which and the points of a region of a 3-dimensional space there is a reciprocal correspondence such that for each point of the space there is a unique set of values for u^1, u^2, u^3, and such that for every set of values for u^1, u^2, u^3, within limits depending upon the boundaries of the region, there exists a corresponding single point of the region. Then the functions u^1, u^2, u^3 can be taken as co-ordinates of the region. Such co-ordinates are called General or Curvilinear Co-ordinates.

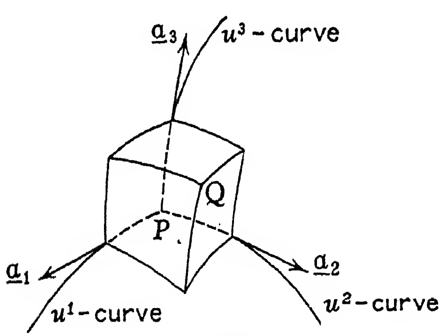


Fig. 36.

At any point $P(u^1, u^2, u^3)$ within such a 3-dimensional region the level surfaces of the functions u^2 and u^3 will intersect in a curve along which the function u^1 alone will vary. This curve is called the u^1-curve. Similarly, there will be a u^2 and a u^3-curve, the intersections of the level surfaces of the functions u^3 and u^1 and

of the functions u^1 and u^2, respectively, along which only u^2 and u^3 respectively will vary. See Fig. 36.

If $\mathbf{r}$ denote the position-vector of the point $P(u^1, u^2, u^3)$ with respect to an arbitrary origin, then $\mathbf{r}$ can be considered as a function of u^1, u^2, u^3, so that:

$$\mathbf{r} = \mathbf{r}(u^1, u^2, u^3).$$

By differentiation:

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u^1} du^1 + \frac{\partial \mathbf{r}}{\partial u^2} du^2 + \frac{\partial \mathbf{r}}{\partial u^3} du^3,$$

or:

$$d\mathbf{r} = \mathfrak{a}_1 du^1 + \mathfrak{a}_2 du^2 + \mathfrak{a}_3 du^3, \tag{1}$$

where:

$$\mathfrak{a}_1 = \frac{\partial \mathbf{r}}{\partial u^1}, \qquad \mathfrak{a}_2 = \frac{\partial \mathbf{r}}{\partial u^2}, \qquad \mathfrak{a}_3 = \frac{\partial \mathbf{r}}{\partial u^3} \tag{2}$$

The vectors $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ are called the Unitary Vectors associated with the point P; they are directed tangentially to the u^1, u^2, u^3-curves respectively in the sense of u^1, u^2, u^3 increasing.

The unitary vectors associated with the point P can be considered as constituting a base-system of reference for all vectors associated with the point P. This base-system will be called the U-system.

If $\mathfrak{a}^1$, $\mathfrak{a}^2$, $\mathfrak{a}^3$ denote the vectors of the system reciprocal to the $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$-system, then, by equations (1), (2), and (3), Art. 19:

$$\mathfrak{a}^1 = \frac{\mathfrak{a}_2 \times \mathfrak{a}_3}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]}, \qquad \mathfrak{a}^2 = \frac{\mathfrak{a}_3 \times \mathfrak{a}_1}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]}, \qquad \mathfrak{a}^3 = \frac{\mathfrak{a}_1 \times \mathfrak{a}_2}{[\mathfrak{a}_1\mathfrak{a}_2\mathfrak{a}_3]}; \tag{3}$$

$$\begin{aligned} \mathfrak{a}^1 \cdot \mathfrak{a}_1 &= \mathfrak{a}^2 \cdot \mathfrak{a}_2 = \mathfrak{a}^3 \cdot \mathfrak{a}_3 = 1, \\ \mathfrak{a}^2 \cdot \mathfrak{a}_1 &= \mathfrak{a}^3 \cdot \mathfrak{a}_2 = \mathfrak{a}^1 \cdot \mathfrak{a}_3 = 0, \\ \mathfrak{a}^3 \cdot \mathfrak{a}_1 &= \mathfrak{a}^1 \cdot \mathfrak{a}_2 = \mathfrak{a}^2 \cdot \mathfrak{a}_3 = 0; \end{aligned} \tag{4}$$

$$\mathfrak{a}_1 = \frac{\mathfrak{a}^2 \times \mathfrak{a}^3}{[\mathfrak{a}^1\mathfrak{a}^2\mathfrak{a}^3]}, \qquad \mathfrak{a}_2 = \frac{\mathfrak{a}^3 \times \mathfrak{a}^1}{[\mathfrak{a}^1\mathfrak{a}^2\mathfrak{a}^3]}, \qquad \mathfrak{a}_3 = \frac{\mathfrak{a}^1 \times \mathfrak{a}^2}{[\mathfrak{a}^1\mathfrak{a}^2\mathfrak{a}^3]}. \tag{5}$$

It will be convenient to call $\mathfrak{a}^1$, $\mathfrak{a}^2$, $\mathfrak{a}^3$ the Reciprocal Unitary Vectors; they constitute a non-coplanar system which may, if desired, be used as well as the unitary system itself as a base-system of reference. This base-system will be called the R-system.

It is our purpose in the present chapter to use the U and R-systems more or less in parallel, in order to bring clearly into evidence their significant relationship, although in so doing we shall be obliged to introduce certain elementary notions relating

to transformation theory, with which the next chapter is more specifically concerned.

Using the $\mathfrak{a}^1$, $\mathfrak{a}^2$, $\mathfrak{a}^3$-system as a base-system, the differential vector $d\mathfrak{r}$ expressed by equation (1) can be expressed in the alternative form:

$$(6) \qquad d\mathfrak{r} = \mathfrak{a}^1 du_1 + \mathfrak{a}^2 du_2 + \mathfrak{a}^3 du_3,$$

where the differentials du_1, du_2, du_3 will not in general, however, be perfect differentials of functions u_1, u_2, u_3 of the general co-ordinates, since they will be linear functions of the differentials of the co-ordinates which may not be integrable.

From equations (1) and (6):

$$\mathfrak{a}^1 du_1 + \mathfrak{a}^2 du_2 + \mathfrak{a}^3 du_3 = \mathfrak{a}_1 du^1 + \mathfrak{a}_2 du^2 + \mathfrak{a}_3 du^3;$$

or:

$$(7) \qquad \sum_{\lambda=1}^{3} \mathfrak{a}^\lambda du_\lambda = \sum_{i=1}^{3} \mathfrak{a}_i du^i.$$

Upon scalar multiplication of both sides of this equation by $\mathfrak{a}_\lambda$ and by $\mathfrak{a}^i$ in turn, and taking into account equations (4), we find:

$$(8) \qquad du_\lambda = \sum_{i=1}^{3} \mathfrak{a}_\lambda \cdot \mathfrak{a}_i du^i, \quad du^i = \sum_{\lambda=1}^{3} \mathfrak{a}^i \cdot \mathfrak{a}^\lambda du_\lambda, \qquad \lambda, i = 1, 2, 3.$$

We shall now find new expressions for the unitary vectors in terms of the reciprocal unitary vectors, and vice versa. From the last three equations we find easily that:

$$\sum_{\lambda=1}^{3} \left(\mathfrak{a}^\lambda - \sum_{i=1}^{3} \mathfrak{a}^i \cdot \mathfrak{a}^\lambda \mathfrak{a}_i \right) du_\lambda = 0,$$

$$\sum_{i=1}^{3} \left(\mathfrak{a}_i - \sum_{\lambda=1}^{3} \mathfrak{a}_\lambda \cdot \mathfrak{a}_i \mathfrak{a}^\lambda \right) du^i = 0.$$

Hence, since the differentials are arbitrary:

$$(9) \qquad \mathfrak{a}^\lambda = \sum_{i=1}^{3} \mathfrak{a}^i \cdot \mathfrak{a}^\lambda \mathfrak{a}_i, \qquad \mathfrak{a}_i = \sum_{\lambda=1}^{3} \mathfrak{a}_\lambda \cdot \mathfrak{a}_i \mathfrak{a}^\lambda.$$

§74

Abbreviated Notation—The Summation Convention

The scalar products of the unitary vectors and the scalar products of the reciprocal unitary vectors which occur in equations

(8) and (9), Art. 73, will often appear. For them we shall, therefore, introduce the abbreviations:

(1) $$g_{\lambda i} = \mathfrak{a}_\lambda \cdot \mathfrak{a}_i = \mathfrak{a}_i \cdot \mathfrak{a}_\lambda = g_{i\lambda};$$

(2) $$g^{i\lambda} = \mathfrak{a}^i \cdot \mathfrak{a}^\lambda = \mathfrak{a}^\lambda \cdot \mathfrak{a}^i = g^{\lambda i}.$$

Furthermore, in order to avoid extensive use of summation signs, we shall frequently use the following Summation Convention:

If any term contain the same letter twice as an identifying index, then summation of such terms for all values of the index is to be understood.

In the present and the next following chapter all literal indices cover the range of values 1, 2, 3.

Equations (8) and (9), Art. 73, can now be written in the abbreviated forms:

(3) $$du_\lambda = g_{\lambda i} du^i, \qquad du^i = g^{i\lambda} du_\lambda;$$

(4) $$\mathfrak{a}^\lambda = g^{\lambda i}\mathfrak{a}_i, \qquad \mathfrak{a}_i = g_{i\lambda}\mathfrak{a}^\lambda.$$

Since the indices λ, i cover the range of values 1, 2, 3, each of these equations is typical of a corresponding set of three equations.

For the determinants (g, g') of the g-coefficients of the sets for which the second and first of equations (4) are typical, we have:

(5) $$\begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix}, \text{ with } g_{ji} = g_{ij};$$

(6) $$\begin{vmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{vmatrix}, \text{ with } g^{\mu\lambda} = g^{\lambda\mu}.$$

If we solve, by the method of determinants, for $\mathfrak{a}^\lambda$ and for $\mathfrak{a}_i$, respectively, the two sets of equations for which the second and first of equations (4) are typical, we find:

(7) $$\mathfrak{a}^\lambda = G^{i\lambda}\mathfrak{a}_i, \qquad \mathfrak{a}_i = G_{\lambda i}\mathfrak{a}^\lambda,$$

where $G^{i\lambda}$ is the cofactor of the element $g_{i\lambda}$ in the i row and λ column of the determinant g divided by g, and where $G_{\lambda i}$ is the cofactor of the element $g^{\lambda i}$ in the λ row and the i column of the determinant g' divided by g'. From equations (4) and (7):

$$g^{\lambda i}\mathfrak{a}_i = G^{i\lambda}\mathfrak{a}_i, \qquad g_{i\lambda}\mathfrak{a}^\lambda = G_{\lambda i}\mathfrak{a}^\lambda.$$

Hence, since the $\mathfrak{a}$'s are non-coplanar, we have:

(9) $$g^{\lambda i} = G^{i\lambda}, \qquad g_{i\lambda} = G_{\lambda i}.$$

Upon scalar multiplication of the second of equations (4) by $\mathfrak{a}^k$, taking account of equations (2), we find:

$$g_{i\lambda}g^{k\lambda} = g_i^k, \tag{10}$$

where $g_i^k = 1$, if $k = i$, and $g_i^k = 0$, if $k \neq i$. For example:

$$g_1^k = g_{11}g^{k1} + g_{12}g^{k2} + g_{13}g^{k3} = 1, \qquad \text{if } k = 1;$$
$$g_1^2 = g_{11}g^{21} + g_{12}g^{22} + g_{13}g^{23} = 0.$$

The significance of g_i^i should be carefully noted:

$$\begin{aligned} g_i^i = g_{i\lambda}g^{i\lambda} = & \quad g_{11}g^{11} + g_{12}g^{12} + g_{13}g^{13} \\ & + g_{21}g^{21} + g_{22}g^{22} + g_{23}g^{23} \\ & + g_{31}g^{31} + g_{32}g^{32} + g_{33}g^{33} \\ = & \quad 1 + 1 + 1 = 3. \end{aligned}$$

It is possible to consider the g-coefficients as forms of operators. Referring to equations (3), for example, it is seen that the typical differential du_λ on the R-system is obtained from the corresponding differentials on the U-system, of which du^i is typical, by causing $g_{\lambda i}$, considered as an operator, to act upon du^i considered as an operand; and, conversely, it is seen that the typical differential du^i on the U-system is obtained from the corresponding differentials on the R-system, of which du_λ is typical, by causing $g^{i\lambda}$, considered as an operator, to act upon du_λ, considered as an operand. In the former case, $g_{\lambda i}$ operates to lower the superscript i of the operand du^i and to change it into λ; in the latter case, $g^{i\lambda}$ operates to raise the subscript λ of the operand du_λ and to change it into i. Similar interpretation can be given to the action of the g's in equations (4).

Using the properties of the g's just mentioned, it is possible to introduce various equivalent forms for the quantities with which we have to deal. We shall consider several cases by way of example.

For the infinitesimal vector $d\mathbf{r}$ we have:

$$\begin{aligned} d\mathbf{r} &= \mathfrak{a}_i du^i \\ &= \mathfrak{a}_i g^{i\lambda} du_\lambda \\ &= \mathfrak{a}^\lambda g_{\lambda i} du^i \\ &= \mathfrak{a}^\lambda du_\lambda. \end{aligned} \tag{11}$$

A fixed vector[1)] $\mathbf{p}$ associated with the point $P(u^1, u^2, u^3)$ can be expressed on the U and R-base-systems as follows:

$$\mathbf{p} = p^i\mathfrak{a}_i = p_\lambda\mathfrak{a}^\lambda, \tag{12}$$

[1)] A fixed vector is one which does not change, as does a base-vector, with change of co-ordinate system.

where the p's are measure-numbers of $\mathbf{p}$. In analogy with equations (3) we have:

$$p_\lambda = g_{\lambda i} p^i, \qquad\qquad p^i = g^{i\lambda} p_\lambda. \tag{13}$$

It follows, in analogy with equations (11), that:

$$\begin{aligned} \mathbf{p} \quad & p^i \mathbf{a}_i \\ &= g^{i\lambda} p_\lambda \mathbf{a}_i \\ &= g_{\lambda i} p^i \mathbf{a}^\lambda \\ &= p_\lambda \mathbf{a}^\lambda. \end{aligned} \tag{14}$$

As another simple example we have in the case of the idemfactor:

$$\begin{aligned} \mathrm{I} &= \mathbf{a}_i \mathbf{a}^i \\ &= g^{i\lambda} \mathbf{a}_i \mathbf{a}_\lambda \\ & \quad g_{i\lambda} \mathbf{a}^\lambda \mathbf{a}^i \\ & \quad \mathbf{a}^\lambda \mathbf{a}_\lambda. \end{aligned} \tag{15}$$

A more complicated example is furnished by a dyadic $\mathbf{\Phi}$, for which we have:

$$\begin{aligned} \mathbf{\Phi} &= a^{ij} \mathbf{a}_i \mathbf{a}_j \\ &= a^{ij} g_{i\lambda} \mathbf{a}^\lambda \mathbf{a}_j \\ &= a^{ij} g_{j\mu} \mathbf{a}_i \mathbf{a}^\mu \\ &= a^{ij} g_{i\lambda} g_{j\mu} \mathbf{a}^\lambda \mathbf{a}^\mu. \end{aligned} \tag{16}$$

For brevity let:

$$\begin{aligned} a_{\lambda\cdot}^{\cdot j} &= a^{ij} g_{i\lambda}, \\ a^{i\cdot}_{\cdot\mu} &= a^{ij} g_{\mu j}, \\ a_{\lambda\mu} &= a^{ij} g_{i\lambda} g_{j\mu}. \end{aligned} \tag{17}$$

We can then write:

$$\begin{aligned} \mathbf{\Phi} &= a^{ij} \mathbf{a}_i \mathbf{a}_j \\ &= a_{\lambda\cdot}^{\cdot j} \mathbf{a}^\lambda \mathbf{a}_j \\ &= a^{i\cdot}_{\cdot\mu} \mathbf{a}_i \mathbf{a}^\mu \\ &= a_{\lambda\mu} \mathbf{a}^\lambda \mathbf{a}^\mu. \end{aligned} \tag{18}$$

§75

Differential Quadratic Forms

If Q denotes a point infinitely near to the point P (see Fig. 36) with the co-ordinates $u^1 + du^1$, $u^2 + du^2$, $u^3 + du^3$, and if ds denote the magnitude of the infinitesimal distance from P to Q, then:

$$\overline{ds}^2 = d\mathbf{r} \cdot d\mathbf{r}, \tag{1}$$

where $d\mathbf{r}$ is the position-vector of Q with respect to P. Using equation (1), Art. 73, we find:

$$\begin{aligned}\overline{ds}^2 = d\mathbf{r}\cdot d\mathbf{r} = \quad & \mathfrak{a}_1\cdot\mathfrak{a}_1du^1du^1 + \mathfrak{a}_1\cdot\mathfrak{a}_2du^1du^2 + \mathfrak{a}_1\cdot\mathfrak{a}_3du^1du^3\\ & +\mathfrak{a}_2\cdot\mathfrak{a}_1du^2du^1 + \mathfrak{a}_2\cdot\mathfrak{a}_2du^2du^2 + \mathfrak{a}_2\cdot\mathfrak{a}_3du^2du^3\\ & +\mathfrak{a}_3\cdot\mathfrak{a}_1du^3du^1 + \mathfrak{a}_3\cdot\mathfrak{a}_2du^3du^2 + \mathfrak{a}_3\cdot\mathfrak{a}_3du^3du^3.\end{aligned}\tag{2}$$

Using equation (6), Art. 73, we obtain the alternative form:

$$\begin{aligned}\overline{ds}^2 = d\mathbf{r}\cdot d\mathbf{r} = \quad & \mathfrak{a}^1\cdot\mathfrak{a}^1du_1du_1 + \mathfrak{a}^1\cdot\mathfrak{a}^2du_1du_2 + \mathfrak{a}^1\cdot\mathfrak{a}^3du_1du_3\\ & +\mathfrak{a}^2\cdot\mathfrak{a}^1du_2du_1 + \mathfrak{a}^2\cdot\mathfrak{a}^2du_2du_2 + \mathfrak{a}^2\cdot\mathfrak{a}^3du_2du_3\\ & +\mathfrak{a}^3\cdot\mathfrak{a}^1du_3du_1 + \mathfrak{a}^3\cdot\mathfrak{a}^2du_3du_2 + \mathfrak{a}^3\cdot\mathfrak{a}^3du_3du_3.\end{aligned}\tag{3}$$

In equation (2) $\overline{ds}^2$ is expressed as a homogeneous quadratic function of the differentials of the co-ordinates themselves, and in equation (3) as a similar function of the non-integrable (in general) differentials du_1, du_2, du_3. The scalar product coefficients of both forms must be considered, in general, as functions of the co-ordinates, since the unitary and reciprocal unitary vectors will vary in general from point to point.

It is convenient to abbreviate the expressions for these differential forms, first by writing, as in Art. 74:

$$g_{ji} = g_{ij} = \mathfrak{a}_i\cdot\mathfrak{a}_j, \qquad g^{\mu\lambda} = g^{\lambda\mu} = \mathfrak{a}^\lambda\cdot\mathfrak{a}^\mu,$$

and further by use of the summation convention[1)] introduced in Art. 74. Equations (2) and (3) respectively can then be written:

$$\overline{ds}^2 = g_{ij}du^idu^j;\tag{4}$$

$$\overline{ds}^2 = g^{\lambda\mu}du_\lambda du_\mu.\tag{5}$$

The g-coefficients in these forms must, of course, be considered in general as functions of the co-ordinates; they will be called Metrical Coefficients, for reasons which will be obvious presently.

The forms (4) and (5) will be called, respectively, the Fundamental Differential Quadratic Form and the Reciprocal Differential Quadratic Form.

An equivalent form in which g's do not appear explicitly can be obtained from the form (4) with the aid of the first of equations (3), Art. 74. It is as follows:

$$\overline{ds}^2 = du^idu_i.\tag{6}$$

[1)] From now on frequent use will be made of this very useful convention; at first the reader may find it desirable to write out in full the expressions which it implies.

§76

Geometrical Significance of the Metrical Coefficients

The values of the metrical coefficients at any point $P(u^1, u^2, u^3)$ must, of course, depend upon the nature of the co-ordinate system which is used.

For the g_{ij}-coefficients with like indices we have, in accordance with their definitions:

$$
\begin{aligned}
g_{11} &= \mathbf{a}_1 \cdot \mathbf{a}_1 = a_1 a_1; \\
g_{22} &= \mathbf{a}_2 \cdot \mathbf{a}_2 = a_2 a_2; \\
g_{33} &= \mathbf{a}_3 \cdot \mathbf{a}_3 = a_3 a_3;
\end{aligned} \tag{1}
$$

where a_1, a_2, a_3, respectively, are the magnitudes of the unitary vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. If $ds_{(1)}$, $ds_{(2)}$, $ds_{(3)}$ denote the magnitudes of $d\mathbf{r}$ when taken in the directions of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, then from equations (2), Art. 75, and equations (1) above we have:

$$
g_{11} = \left(\frac{du^1}{ds_{(1)}}\right)^{-2}; \qquad g_{22} = \left(\frac{du^2}{ds_{(2)}}\right)^{-2}; \qquad g_{33} = \left(\frac{du_3}{ds_{(3)}}\right)^{-2} \tag{2}
$$

Hence, g_{11}, g_{22}, g_{33} are respectively equal to the squares of the reciprocals of the space rates of increase of the co-ordinates in the directions of the corresponding unitary vectors.

For the g_{ij} coefficients with unlike indices we have:

$$
\begin{aligned}
g_{32} &= g_{23} = \mathbf{a}_2 \cdot \mathbf{a}_3 = a_2 a_3 \cos(\mathbf{a}_2, \mathbf{a}_3); \\
g_{13} &= g_{31} = \mathbf{a}_3 \cdot \mathbf{a}_1 = a_3 a_1 \cos(\mathbf{a}_3, \mathbf{a}_1); \\
g_{21} &= g_{12} = \mathbf{a}_1 \cdot \mathbf{a}_2 = a_1 a_2 \cos(\mathbf{a}_1, \mathbf{a}_2).
\end{aligned}
$$

Hence, upon taking account of equations (1):

$$
\begin{aligned}
g_{32} &= g_{23} = \sqrt{g_{22}g_{33}} \cos(\mathbf{a}_2, \mathbf{a}_3); \\
g_{13} &= g_{31} = \sqrt{g_{33}g_{11}} \cos(\mathbf{a}_3, \mathbf{a}_1); \\
g_{21} &= g_{12} = \sqrt{g_{11}g_{22}} \cos(\mathbf{a}_1, \mathbf{a}_2);
\end{aligned} \tag{3}
$$

where the positive signs for the radicals are taken, since a_1, a_2, a_3 are intrinsically positive.

From equations (3) we have for the cosines of the angles between the unitary vectors in pairs:

$$
\begin{aligned}
\cos(\mathbf{a}_2, \mathbf{a}_3) &= \frac{g_{23}}{\quad}. \\
\cos(\mathbf{a}_3, \mathbf{a}_1) &= \frac{\quad}{\sqrt{g_{33}g_{11}}}, \\
\cos(\mathbf{a}_1, \mathbf{a}_2) &= \frac{g_{12}}{\sqrt{g_{11}g_{22}}}.
\end{aligned} \tag{4}
$$

The cosine of the angle between any two infinitesimal vectors:

$$\text{(5)} \qquad d\mathbf{s} = \mathbf{a}_1 du^1 + \mathbf{a}_2 du^2 + \mathbf{a}_3 du^3 = \mathbf{a}_i du^i,$$

$$\text{(6)} \qquad \delta\mathbf{s} = \mathbf{a}_1 \delta u^1 + \mathbf{a}_2 \delta u^2 + \mathbf{a}_3 \delta u^3 = \mathbf{a}_j \delta u^j,$$

is obtained by forming the scalar product of $d\mathbf{s}$ and $\delta\mathbf{s}$. Upon performing this operation we obtain the bilinear differential form:

$$\text{(7)} \qquad \begin{aligned} d\mathbf{s} \cdot \delta\mathbf{s} = ds\delta s \cos(d\mathbf{s}, \delta\mathbf{s}) &= \mathbf{a}_i \cdot \mathbf{a}_j du^i \delta u^j \\ &= g_{ij} du^i \delta u^j, \end{aligned}$$

where ds, δs denote the magnitudes of $d\mathbf{s}$, $\delta\mathbf{s}$, and where summation is implied by the summation convention for both the indices i and j over the range of values 1, 2, 3. We therefore have:

$$\text{(8)} \qquad \cos(d\mathbf{s} \cdot \delta\mathbf{s}) = g_{ij} \frac{du^i}{ds} \frac{\delta u^j}{\delta s}.$$

The significance of the coefficients in the reciprocal differential quadratic form (5), Art. 75, can be obtained in a similar way to that followed above in discussing the coefficients of the corresponding form (4), Art. 75. In fact, we can at once write down the following equations by analogy:

$$\text{(9)} \qquad \begin{aligned} g^{11} &= \mathbf{a}^1 \cdot \mathbf{a}^1 = a^1 a^1; \\ g^{22} &= \mathbf{a}^2 \cdot \mathbf{a}^2 = a^2 a^2; \\ g^{33} &= \mathbf{a}^3 \cdot \mathbf{a}^3 = a^3 a^3; \end{aligned}$$

$$\text{(10)} \qquad {}^{-2}; \qquad g^{22} = \left(\frac{du_2}{ds^{(2)}}\right)^{-2}; \qquad g^{33} = \left(\frac{du_3}{ds^{(3)}}\right.$$

$$\text{(11)} \qquad \begin{aligned} g^{32} = g^{23} &= \sqrt{g^{22}g^{33}} \cos(\mathbf{a}^2, \mathbf{a}^3); \\ g^{13} = g^{31} &= \sqrt{g^{33}g^{11}} \cos(\mathbf{a}^3, \mathbf{a}^1); \\ g^{21} = g^{12} &= \sqrt{g^{11}g^{22}} \cos(\mathbf{a}^1, \mathbf{a}^2); \end{aligned}$$

$$\text{(12)} \qquad \begin{aligned} \cos(\mathbf{a}^2, \mathbf{a}^3) &= \frac{g^{23}}{\sqrt{g^{22}g^{33}}}; \\ \cos(\mathbf{a}^3, \mathbf{a}^1) &= \frac{g}{\sqrt{g^{33}g^{11}}}; \\ \cos(\mathbf{a}^1, \mathbf{a}^2) &= \frac{g^{12}}{\sqrt{g^{11}g^{22}}}; \end{aligned}$$

$$\text{(13)} \qquad d\mathbf{s} = \mathbf{a}^1 du_1 + \mathbf{a}^2 du_2 + \mathbf{a}^3 du_3 = \mathbf{a}^\lambda du_\lambda;$$

$$\text{(14)} \qquad \delta\mathbf{s} = \mathbf{a}^1 \delta u_1 + \mathbf{a}^2 \delta u_2 + \mathbf{a}^3 \delta u_3 = \mathbf{a}^\mu \delta u_\mu;$$

$$\text{(15)} \qquad \begin{aligned} d\mathbf{s} \cdot \delta\mathbf{s} = ds\delta s \cos(d\mathbf{s} \cdot \delta\mathbf{s}) &= \mathbf{a}^\lambda \cdot \mathbf{a}^\mu du_\lambda \delta u_\mu \\ &= g^{\lambda\mu} du_\lambda \delta u_\mu; \end{aligned}$$

$$\text{(16)} \qquad \cos(d\mathbf{s}, \delta\mathbf{s}) = g^{\lambda\mu} \frac{du_\lambda}{ds} \frac{\delta u_\mu}{\delta s}.$$

§77

Line, Surface, and Volume Elements

If $d\mathbf{s}_{(1)}$, $d\mathbf{s}_{(2)}$, $d\mathbf{s}_{(3)}$ denote the infinitesimal vectors at the point $P(u^1, u^2, u^3)$ obtained by taking in turn the infinitesimal vector $d\mathbf{s}$ in the directions of $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$, then:

$$(1) \qquad \begin{aligned} d\mathbf{s}_{(1)} &= \mathfrak{a}_1 du^1; \\ d\mathbf{s}_{(2)} &= \mathfrak{a}_2 du^2; \\ d\mathbf{s}_{(3)} &= \mathfrak{a}_3 du^3; \end{aligned}$$

$$(2) \qquad \begin{aligned} ds_{(1)} &= \sqrt{g_{11}} du^1; \\ ds_{(2)} &= \sqrt{g_{22}} du^2; \\ ds_{(3)} &= \sqrt{g_{33}} du^3. \end{aligned}$$

For the magnitude $d\sigma_{(1)}$ of the area of the parallelogram constructed upon infinitesimal line-vectors representing $d\mathbf{s}_{(2)}$ and $d\mathbf{s}_{(3)}$ we have:

$$\begin{aligned} d\sigma_{(1)} = |d\mathbf{s}_{(2)} \times d\mathbf{s}_{(3)}| &= \pm \mathfrak{a}_2 \times \mathfrak{a}_3 du^2 du^3 \\ &= \pm a_2 a_3 \sin(\mathfrak{a}_2, \mathfrak{a}_3) du^2 du^3 \\ &= \sqrt{g_{22} g_{33} - g_{23}^2}\, du^2 du^3, \end{aligned}$$

by equations (1) and (3), Art. 76; and by cyclical permutation of the subscripts and superscripts we can obtain the expressions for the corresponding magnitudes $d\sigma_{(2)}$ and $d\sigma_{(3)}$. Hence:

$$(3) \qquad \begin{aligned} d\sigma_{(1)} &= \sqrt{g_{22} g_{33} - g_{23}^2}\, du^2 du^3; \\ d\sigma_{(2)} &= \sqrt{g_{33} g_{11} - g_{31}^2}\, du^3 du^1; \\ d\sigma_{(3)} &= \sqrt{g_{11} g_{22} - g_{12}^2}\, du^1 du^2. \end{aligned}$$

If $d\tau$ denote the magnitude of the volume of an elementary parallelepiped constructed upon the infinitesimal line-vectors representing $d\mathbf{s}_{(1)}$, $d\mathbf{s}_{(2)}$, $d\mathbf{s}_{(3)}$, then:

$$d\tau = \pm \mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3\, du^1 du^2 du^3;$$

but, with the aid of equations (1′), Art. 20:

$$\mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3 = \mathfrak{a}_1 \cdot [(\mathfrak{a}^1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\, \mathfrak{a}_1 + (\mathfrak{a}^2 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\, \mathfrak{a}_2 + (\mathfrak{a}^3 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\, \mathfrak{a}_3]$$

$$= \mathfrak{a}_1 \cdot \frac{[(\mathfrak{a}_2 \times \mathfrak{a}_3 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\, \mathfrak{a}_1 + (\mathfrak{a}_3 \times \mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\, \mathfrak{a}_2 + (\mathfrak{a}_1 \times \mathfrak{a}_2 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)\, \mathfrak{a}_3]}{\mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3},$$

with the aid of equations (1), Art. 19; hence:

$$
\begin{aligned}
(\mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3)^2 = \; & (\mathfrak{a}_1 \cdot \mathfrak{a}_1)\,[(\mathfrak{a}_2 \cdot \mathfrak{a}_2)\,(\mathfrak{a}_3 \cdot \mathfrak{a}_3) - (\mathfrak{a}_2 \cdot \mathfrak{a}_3)\,(\mathfrak{a}_3 \cdot \mathfrak{a}_2)] \\
& + (\mathfrak{a}_1 \cdot \mathfrak{a}_2)\,[(\mathfrak{a}_2 \cdot \mathfrak{a}_3)\,(\mathfrak{a}_3 \cdot \mathfrak{a}_1) - (\mathfrak{a}_2 \cdot \mathfrak{a}_1)\,(\mathfrak{a}_3 \cdot \mathfrak{a}_3)] \\
& + (\mathfrak{a}_1 \cdot \mathfrak{a}_3)\,[(\mathfrak{a}_2 \cdot \mathfrak{a}_1)\,(\mathfrak{a}_3 \cdot \mathfrak{a}_2) - (\mathfrak{a}_2 \cdot \mathfrak{a}_2)\,(\mathfrak{a}_3 \cdot \mathfrak{a}_1)] \\
= \; & g_{11}[g_{22}g_{33} - g_{23}g_{32}] \\
& + g_{12}[g_{23}g_{31} - g_{21}g_{33}] \\
& + g_{13}[g_{21}g_{32} - g_{22}g_{31}];
\end{aligned}
$$

therefore:

$$
(\mathfrak{a}_1 \cdot \mathfrak{a}_2 \times \mathfrak{a}_3) = \pm \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & \\ g_{31} & g_{32} & \end{vmatrix}^{\frac{1}{2}} \pm \sqrt{g}.
$$

Upon substituting this result in the above equation for $d\tau$, we find:

$$
(4) \qquad d\tau = \sqrt{g}\,du^1du^2du^3.
$$

When the R-base-system is used, corresponding to equations (1) to (4) inclusive, we have:

$$
(5) \qquad
\begin{aligned}
d\mathbf{s}^{(1)} &= \mathfrak{a}^1 du_1, \\
d\mathbf{s}^{(2)} &= \mathfrak{a}^2 du_2, \\
d\mathbf{s}^{(3)} &= \mathfrak{a}^3 du_3;
\end{aligned}
$$

$$
(6) \qquad
\begin{aligned}
ds^{(1)} &= \sqrt{g^{11}}\,du_1, \\
ds^{(2)} &= \sqrt{g^{22}}\,du_2, \\
ds^{(3)} &= \sqrt{g^{33}}\,du_3;
\end{aligned}
$$

$$
(7) \qquad
\begin{aligned}
d\sigma^{(1)} &= \sqrt{g^{22}g^{33} - (g^{23})^2}\,du_2du_3, \\
d\sigma^{(2)} &= \sqrt{g^{33}g^{11} - (g^{31})^2}\,du_3du_1, \\
d\sigma^{(3)} &= \sqrt{g^{11}g^{22} - (g^{12})^2}\,du^1du^2;
\end{aligned}
$$

$$
(8) \qquad d\tau' = \sqrt{g'}\,du_1du_2du_3.
$$

The element of volume whose magnitude is denoted by $d\tau$ is, if $g \neq 0$, equivalent to the element of volume bounded by portions of level surfaces of the co-ordinates u^1, u^2, u^3 passing through the points $P(u^1, u^2, u^3)$ and $Q(u^1 + du^1, u^2 + du^2, u^3 + du^3)$, if differences in volume which are infinitely small in comparison with the elements of volume in question are ignored. Similarly, the corresponding edges and surfaces of the two elements can be considered as equivalent, if differences which are infinitesimal in comparison with the quantities themselves are neglected. The proof of these statements is not difficult to supply, and will be omitted here.

§78
Determination of Metrical Coefficients
The Dummy Index Rule

The results found in Arts. 76 and 77 show that the metrical properties of our 3-dimensional space can be expressed with respect to any general co-ordinate system in terms of the g_{ij}-coefficients of the differential quadratic form (4), Art. 75, or in terms of the corresponding $g^{\lambda\mu}$-coefficients of the corresponding differential quadratic form (5), Art. 75. In other words, when the g_{ij} or the $g^{\lambda\mu}$-coefficients have been determined for any system of general co-ordinates, lengths of lines, angles between any two directions, areas, and volumes can be expressed in terms of them.

With the introduction of a general system of co-ordinates, the determination of the metrical coefficients of its differential quadratic forms is, therefore, a matter of importance.

In some cases the differential quadratic forms can at once be written down from geometrical considerations, and the metrical coefficients determined by inspection. This is certainly the case for a rectangular Cartesian system of co-ordinates (x^1, x^2, x^3), for which we have the differential quadratic form:

$$\bar{ds}^2 = dx^1dx^1 + dx^2dx^2 + dx^3dx^3. \tag{1}$$

From this we infer that the metrical coefficients g_{11}, g_{22}, g_{33} are all equal to unity, and that the coefficients g_{23}, g_{31}, g_{12} are all equal to zero. If, now, any system of general co-ordinates (u^1, u^2, u^3) be introduced, the corresponding differential quadratic form must be derivable from the form (1) by a process of transformation, and from the form thus obtained the metrical coefficients for the general system can be found when the transformation equations are given, as will now be shown.

Let us suppose that the x and the u-co-ordinates are related by the set of equations:

$$\begin{aligned} x^1 &= x^1(u^1, u^2, u^3); \\ x^2 &= x^2(u^1, u^2, u^3); \\ x^3 &= x^3(u^1, u^2, u^3), \end{aligned} \tag{2}$$

where u^1, u^2, u^3 are supposed mutually independent.

Upon differentiation of these equations we obtain the following transformation equations for the differentials of the co-ordinates:

$$
\begin{aligned}
dx^1 &= \frac{\partial x^1}{\partial u^1}\,du^1 + \frac{\partial x^1}{\partial u^2}\,du^2 + \frac{\partial x^1}{\partial u^3}\,du^3,\\
(3) \qquad dx^2 &= \frac{\partial x^2}{\partial u^1}\,du^1 + \frac{\partial x^2}{\partial u^2}\,du^2 + \frac{\partial x^2}{\partial u^3}\,du^3,\\
dx^3 &= \frac{\partial x^3}{\partial u^1}\,du^1 + \frac{\partial x^3}{\partial u^2}\,du^2 + \frac{\partial x^3}{\partial u^3}\,du^3.
\end{aligned}
$$

The determinant

$$
\begin{vmatrix}
\dfrac{\partial x^1}{\partial u^1} & \dfrac{\partial x^1}{\partial u^2} & \dfrac{\partial x^1}{\partial u^3}\\
\dfrac{\partial x^2}{\partial u^1} & \dfrac{\partial x^2}{\partial u^2} & \dfrac{\partial x^2}{\partial u^3}\\
\dfrac{\partial x^3}{\partial u^1} & \dfrac{\partial x^3}{\partial u^2} & \dfrac{\partial x^3}{\partial u^3}
\end{vmatrix}
$$

is called the Jacobian of the transformation, and it cannot vanish identically, since we assume the independence of u^1, u^2, u^3.

Making use of the summation convention, introduced in Art. 74, we can write equations (1) and (3) in the abbreviated forms:

$$(4) \qquad \overline{ds}^2 = dx^k dx^k, \qquad (k = 1, 2, 3);$$

$$(5) \qquad dx^k = \frac{\partial x^k}{\partial u^i}\,du^i, \qquad (i, k = 1, 2, 3).$$

Corresponding to the form (4) for the x-differentials, we have, for the u-differentials:

$$(6) \qquad \overline{ds}^2 = g_{ij}\,du^i du^j, \qquad (i, j = 1, 2, 3).$$

If, in attempting to transform the right-hand member of equation (4) into a quadratic function of the du's equivalent to the form (6), we should substitute the value for dx^k given by equation (5), we should obtain an expression involving the index i four times, and confusion would arise in using the summation convention. But this can be avoided by observing that:

$$\frac{\partial x^k}{\partial u^i}\,du^i = \frac{\partial x^k}{\partial u^j}\,du^j,$$

since the values to be assigned to both i and j are 1, 2, 3. We have here an example of the very useful Dummy Index Rule:

In any term in which the summation convention is operative, any letter which appears twice as an index (indicating summation over a given range of values) may be replaced by any other letter, not already in use as an index, with the understanding that it indicates summation over the same range of values.

Accordingly, we can write:

$$\overline{ds}^2 = \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j} du^j du^i, \qquad (i, j, k = 1, 2, 3). \tag{7}$$

Upon comparison of this expression for $\overline{ds}^2$ with that given by equation (6), noting that the du's may be considered as arbitrary, we get:

$$\begin{aligned} g_{ij} &= \frac{\partial x^k}{\partial u^i}\frac{\partial x^k}{\partial u^j}, \\ &= \frac{\partial x^1}{\partial u^i}\frac{\partial x^1}{\partial u^j} + \frac{\partial x^2}{\partial u^i}\frac{\partial x^2}{\partial u^j} + \frac{\partial x^3}{\partial u^i}\frac{\partial x^3}{\partial u^j}, \qquad (i, j = 1, 2, 3). \end{aligned} \tag{8}$$

Having found with the aid of this formula the values of the metrical coefficients for a given U-system, the values of the corresponding coefficients on the corresponding R-system can be found with the aid of the first of equations (9), Art. 74.

We shall exemplify the use of formula (8) by determining the coefficients for several special cases.

Cylindrical co-ordinates. We suppose that:

$$\begin{aligned} u^1 &= \rho, \\ u^2 &= \phi, \\ u^3 &= z, \end{aligned}$$

where ρ, ϕ, z are cylindrical co-ordinates of any point P. See Fig. 37.

The level surfaces of ρ are cylinders coaxial with the z-axis; those of ϕ are meridian planes through the z-axis; and those of z are planes perpendicular to the z-axis.

Choose rectangular Cartesian co-ordinates x^1, x^2, x^3 for the point P such that:

$$\begin{aligned} x^1 &= \rho \cos \phi, \\ x^2 &= \rho \sin \phi, \\ x^3 &= z. \end{aligned} \tag{9}$$

These are the forms for the transformation equations (2) for the case of cylindrical co-ordinates.

Upon inserting these values for x^1, x^2, x^3 in equation (8), we find the following values for the g_{ij} coefficients of the fundamental differential quadratic form for a cylindrical co-ordinate system:

$$\begin{aligned} &g_{11} = 1, &g_{22} = \rho^2, \qquad &g_{33} = 1, \\ g_{32} = {}&g_{23} = 0, \qquad g_{13} = {}&g_{31} = 0, \qquad g_{21} = {}&g_{12} = 0. \end{aligned} \tag{10}$$

The fundamental differential quadratic form for cylindrical co-ordinates will therefore be:

(11) $$\overline{ds}^2 = \overline{d\rho}^2 + \rho^2\overline{d\phi}^2 + \overline{dz}^2.$$

Using the values furnished by equations (10), we find, with the aid of the first of equations (9), Art. 74:

(12) $$g^{11} = 1, \qquad g^{22} = \frac{1}{\rho^2}, \qquad g^{33} = 1,$$
$$g^{32} = g^{23} = 0, \qquad g^{13} = g^{31} = 0, \qquad g^{21} = g^{12} = 0.$$

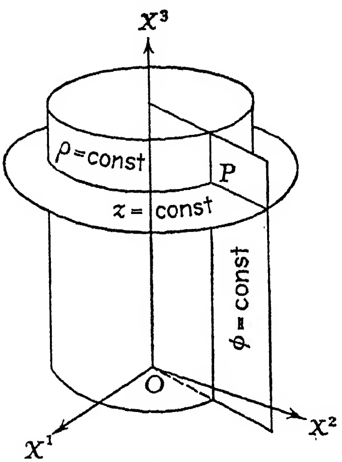


Fig. 37.

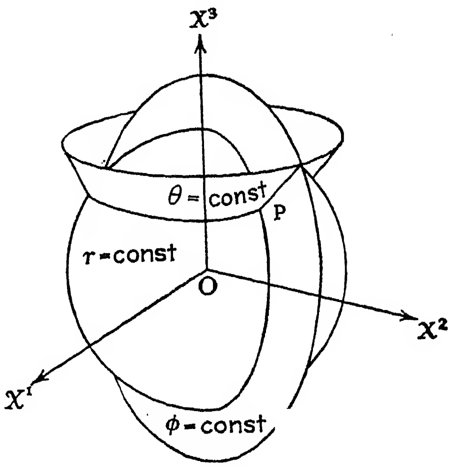


Fig. 38.

Spherical co-ordinates. We next suppose that:

$$u^1 = r,$$
$$u^2 = \theta,$$
$$u^3 = \phi,$$

where r, θ, ϕ are spherical co-ordinates of any point P. See Fig. 38.

The level surfaces of r are spheres with the origin as center; those of θ are cones with the origin as vertex; and those of ϕ are meridian planes.

Choose rectangular Cartesian co-ordinates x^1, x^2, x^3 for the point P such that:

(13) $$x^1 = r \sin\theta \cos\phi,$$
$$x^2 = r \sin\theta \sin\phi,$$
$$x^3 = r \cos\theta.$$

These are the forms for the transformation equations (2) for the case of spherical co-ordinates.

Equation (8) furnishes, with the aid of these equations, the following values for the g_{ij} coefficients of the differential quadratic form for a spherical co-ordinate system:

$$(14)\qquad \begin{matrix} g_{11} = 1; & g_{22} = r^2; & g_{33} = r^2 \sin^2 \theta; \\ g_{32} = g_{23} = 0; & g_{13} = g_{31} = 0; & g_{21} = g_{12} = 0. \end{matrix}$$

The fundamental differential quadratic form for spherical co-ordinates will therefore be:

$$(15)\qquad \overline{ds}^2 = \overline{dr}^2 + r^2\overline{d\theta}^2 + r^2 \sin^2 \theta \overline{d\phi}^2 .$$

From the first of equations (9), Art. 74, and equations (14), we find:

$$(16)\qquad \begin{matrix} g^{11} = 1; & g^{22} = \dfrac{1}{r^2}; & g^{33} = \dfrac{1}{r^2 \sin^2 \theta} \\ g^{32} = g^{23} = 0; & g^{13} = g^{31} = 0; & g^{21} = g^{12} = 0. \end{matrix}$$

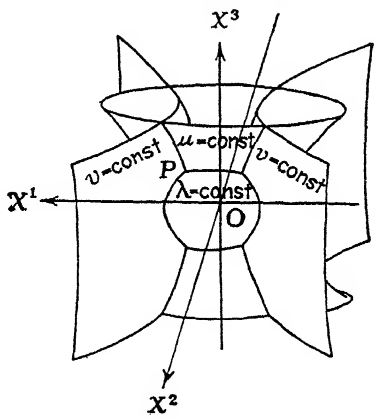


Fig. 39.

Ellipsoidal co-ordinates. As a final example let us suppose that:

$$u^1 = \lambda,$$
$$u^2 = \mu,$$
$$u^3 = \nu,$$

where λ, μ, ν are ellipsoidal co-ordinates of any point P. See Fig. 39.

The level surfaces are ellipsoids, hyperboloids of one sheet, and hyperboloids of two sheets, all confocal with the ellipsoid whose equation in rectangular Cartesian co-ordinates, x^1, x^2, x^3, can be written:

$$(17)\qquad \frac{(x^1)^2}{a^2} + \frac{(x^2)^2}{b^2} + \frac{(x^3)^2}{c^2} = 1, \quad \text{with } a > b > c,$$

where a, b, c denote the magnitudes of semi-principal axes.

The equations of an ellipsoid, hyperboloid of one sheet, and hyperboloid of two sheets confocal with the ellipsoid in question can be written respectively as follows:

$$
\begin{aligned}
&\frac{(x^1)^2}{a^2+\lambda}+\frac{(x^2)^2}{b^2+\lambda}+\frac{(x^3)^2}{c^2+\lambda}=1, && (\lambda > -c^2);\\
(18)\qquad &\frac{(x^1)^2}{a^2+\mu}+\frac{(x^2)^2}{b^2+\mu}+\frac{(x^3)^2}{c^2+\mu}=1, && (-c^2 > \mu > -b^2);\\
&\frac{(x^1)^2}{a^2+\nu}+\frac{(x^2)^2}{b^2+\nu}+\frac{(x^3)^2}{c^2+\nu}=1, && (-b^2 > \nu > -a^2).
\end{aligned}
$$

Through each point of space there will pass one surface of each sort, and to each point there will correspond a unique set of values for λ, μ, ν which can therefore be taken as co-ordinates of the point.

By solving the last three equations for x^1, x^2, x^3, we find:

$$
\begin{aligned}
&\pm\left[\frac{(a^2+\lambda)(a^2+\mu)(a^2+\nu)}{(a^2-b^2)(a^2-c^2)}\right]^{\frac{1}{2}},\\
(19)\qquad x^2 &= \pm\left[\frac{(b^2+\lambda)(b^2+\mu)(b^2+\nu)}{(b^2-c^2)(b^2-a^2)}\right]^{\frac{1}{2}},\\
x^3 &= \pm\left[\frac{(c^2+\lambda)(c^2+\mu)(c^2+\nu)}{(c^2-a^2)(c^2-b^2)}\right]^{\frac{1}{2}}.
\end{aligned}
$$

These are the forms for the transformation equation (2) for the case of ellipsoidal co-ordinates. From these three equations the following equation, which will be useful presently, is easily found:

$$
(20)\quad \frac{(x^1)^2}{(a^2+\lambda)^2}+\frac{(x^2)^2}{(b^2+\lambda)^2}+\frac{(x^3)^2}{(c^2+\lambda)^2}=\frac{(\lambda-\mu)(\lambda-\nu)}{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)},
$$

and two others of the same type, by cyclical permutation of λ, μ, ν.

With the aid of equations (19) we find from equation (8):

$$
\begin{aligned}
g_{11} &= \frac{\partial x^1}{\partial\lambda}\frac{\partial x^1}{\partial\lambda}+\frac{\partial x^2}{\partial\lambda}\frac{\partial x^2}{\partial\lambda}+\frac{\partial x^3}{\partial\lambda}\frac{\partial x^3}{\partial\lambda}\\
&= \frac{1}{4}\left[\frac{(a^2+\mu)(a^2+\nu)}{(a^2+\lambda)(a^2-b^2)(a^2-c^2)}\right]\\
&\quad+\frac{1}{4}\left[\frac{(b^2+\mu)(b^2+\nu}{(b^2+\lambda)(b^2-c^2)(b^2-a^2)}\right]\\
&\quad+\frac{1}{4}\left[\frac{(c^2+\mu)(c^2+\nu)}{(c^2+\lambda)(c^2-a^2)(c^2-b^2)}\right]\\
&= \frac{1}{4}\left[\frac{(x^1)^2}{(a^2+\lambda)^2}+\frac{(x^2)^2}{(b^2+\lambda)^2}+\frac{(x^3)^3}{(c^2+\lambda)^2}\right],
\end{aligned}
$$

or, taking account of equation (20):

$$g_{11} = \frac{1}{4}\frac{(\lambda - \mu)(\lambda - \nu)}{(a^2 + \lambda)(b^2 + \lambda)(c^2 + \lambda)}.$$

The determination of g_{12}, with the aid of equation (8), can be effected as follows:

$$\begin{aligned} g_{12} &= \frac{\partial x^1}{\partial \lambda}\frac{\partial x^1}{\partial \mu} + \frac{\partial x^2}{\partial \lambda}\frac{\partial x^2}{\partial \mu} + \frac{\partial x^3}{\partial \lambda}\frac{\partial x^3}{\partial \mu} \\ &= \frac{1}{4}\frac{(a^2+\nu)(b^2-c^2) + (b^2+\nu)(c^2-a^2) + (c^2+\nu)(a^2-b^2)}{(a^2-b^2)(b^2-c^2)(c^2-a^2)} \\ &= \frac{1}{4}\frac{a^2b^2 - a^2c^2 + b^2c^2 - b^2a^2 + c^2a^2 - c^2b^2}{(a^2-b^2)(b^2-c^2)(c^2-a^2)} \\ &\quad + \frac{1}{4}\frac{\nu(b^2 - c^2 + c^2 - a^2 + a^2 - b^2)}{(a^2-b^2)(b^2-c^2)(c^2-a^2)} \\ &= 0. \end{aligned}$$

In like manner g_{23} and g_{31} can be shown to vanish.

The coefficients g_{22} and g_{33} can be found from the expression for g_{11} found above by cyclical permutation of λ, μ, ν. Hence, for the g_{ij}-coefficients of an ellipsoidal co-ordinate system we have:

$$(21)\qquad \begin{aligned} g_{11} &= \frac{1}{4}\frac{(\lambda-\mu)(\lambda-\nu)}{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)}, & g_{32} &= g_{23} = 0, \\ g_{22} &= \frac{1}{4}\frac{(\mu-\nu)(\mu-\lambda)}{(a^2+\mu)(b^2+\mu)(c^2+\mu)}, & g_{13} &= g_{31} = 0, \\ g_{33} &= \frac{1}{4}\frac{(\nu-\lambda)(\nu-\mu)}{(a^2+\nu)(b^2+\nu)(c^2+\nu)}, & g_{21} &= g_{12} = 0. \end{aligned}$$

Since $g_{12} = g_{23} = g_{31} = 0$, the level surfaces of λ, μ, ν, must intersect orthogonally.

The fundamental differential quadratic form for the ellipsoidal system of co-ordinates will therefore be:

$$(22)\qquad \begin{aligned} \overline{ds}^2 = &\frac{1}{4}\frac{(\lambda-\mu)(\lambda-\nu)}{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)}\overline{d\lambda}^2 \\ &+\frac{1}{4}\frac{(\mu-\nu)(\mu-\lambda)}{(a^2+\mu)(b^2+\mu)(c^2+\mu)}\overline{d\mu}^2 \\ &+\frac{1}{4}\frac{(\nu-\lambda)(\nu-\mu)}{(a^2+\nu)(b^2+\nu)(c^2+\nu)}\overline{d\nu}^2. \end{aligned}$$

Equations (21) and the first of equations (9), Art. 74, give:

$$\begin{aligned} g^{11} &= 4\frac{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)}{(\lambda-\mu)(\lambda-\nu)}, & g^{32} &= g^{23} = 0, \\ g^{22} &= 4\frac{(a^2+\mu)(b^2+\mu)(c^2+\mu)}{(\mu-\nu)(\mu-\lambda)}, & g^{13} &= g^{31} = 0, \\ g^{33} &= 4\frac{(a^2+\nu)(b^2+\nu)(c^2+\nu)}{(\nu-\lambda)(\nu-\mu)}, & g^{21} &= g^{12} = 0. \end{aligned}$$

Having found the values of the g-coefficients for cylindrical, spherical, and ellipsoidal co-ordinate systems, the values of line, surface, and volume elements on these systems can be obtained directly from the equations of Art. 77.

§79

Metrical Coefficients for a Surface

So far in the present chapter we have discussed only co-ordinate systems which function in our 3-dimensional Euclidean space. We have seen that the metrical properties of this space can be expressed in terms of the metrical g-coefficients of the differential quadratic forms appropriate to any system of co-ordinates which we may desire to use.

Let us now consider a system of Gaussian co-ordinates for a surface. Such co-ordinates were defined and used in Art. 31 in connection with the theory of surfaces. It may be recalled that, if u^1, u^2 denote the Gaussian co-ordinates of a point on a surface for which the three rectangular Cartesian co-ordinates are x^1, x^2, x^3, then the equations:

$$x^1 = x^1(u^1, u^2),$$
$$x^2 = x^2(u^1, u^2),$$
$$x^3 = x^3(u^1, u^2),$$

represent the surface considered as a section of the Euclidean 3-dimensional space in which the surface is embedded, provided any two of these three equations can be solved for the parameters u^1, u^2, and that the values thus found for these parameters when substituted in the third equation shall give the rectangular Cartesian equation of the surface.

As was shown in Art. 31, the fundamental differential quadratic form for the surface is as follows:

$$\overline{ds}^2 = g_{11}du^1du^2 + 2g_{12}du^1du^2 + g_{22}du^2du^2.$$

It was also shown in Art. 31 that the metrical properties of the surface could be expressed in terms of the g-coefficients of this form. Furthermore, as was first shown by Gauss, these coefficients can be expressed in terms of quantities which do not involve any quantity which is alien to the surface itself. When this is done the result is a surface geometry which is purely intrinsic.

The analogy of the metrical properties of a surface, as expressed through the metrical coefficients of its fundamental differential quadratic form, with the corresponding metrical properties of our 3-dimensional Euclidean space, as expressed through the metrical coefficients of the fundamental differential quadratic form associated with any general system of co-ordinates, is evident. Nevertheless, there exist certain generic distinctions between the two cases, which are of important significance, and to which attention should be drawn.

In the case of our 3-dimensional Euclidean space it is always possible by change of co-ordinates to reduce the fundamental differential quadratic form to a sum of squares of differentials of the co-ordinates. This is a characteristic of the metrics of Euclidean space.

In the case of a surface, unless its Gaussian curvature, defined in Art. 30, vanishes, it is not possible to reduce its fundamental differential quadratic form to a sum of squares of the differentials of its Gaussian co-ordinates. Hence, a surface for which the Gaussian curvature does not vanish is called a non-Euclidean 2-dimensional space.

The Gaussian curvature does vanish for a flat surface, of course, and also for a developable surface, such as that of a cylinder or of a cone; for these surfaces can be rolled out on a flat surface without stretching and, therefore, without changing their metrics, since the g-coefficient of the corresponding fundamental differential quadratic forms in terms of which, as we have seen in Art. 31, the metrical properties of a surface can be expressed, will not thereby be changed.

In the case of Euclidean space, since it is always possible by transformation of co-ordinates to express its differential quadratic form as a sum of squares of differentials of the co-ordinates, it is possible to reduce each of the g-coefficients in the principal diagonal of their determinant to the value unity, and each of the others to the value zero. In non-Euclidean space this is not possible.

EXERCISES ON CHAPTER VII

1. If $\mathbf{r}$ is the position-vector of a generic point $P(u^1, u^2, u^3)$, show that the following relations are true:

$$\frac{\partial \mathbf{r}}{\partial u^1} \cdot \nabla u^1 = 1, \qquad \frac{\partial \mathbf{r}}{\partial u^1} \cdot \nabla u^2 = 0, \qquad \frac{\partial \mathbf{r}}{\partial u^1} \cdot \nabla u^3 = 0,$$

$$\frac{\partial \mathbf{r}}{\partial u^2} \cdot \nabla u^2 = 1, \qquad \frac{\partial \mathbf{r}}{\partial u^2} \cdot \nabla u^3 = 0, \qquad \frac{\partial \mathbf{r}}{\partial u^2} \cdot \nabla u^1 = 0,$$

$$\frac{\partial \mathbf{r}}{\partial u^3} \cdot \nabla u^3 = 1, \qquad \frac{\partial \mathbf{r}}{\partial u^3} \cdot \nabla u^1 = 0, \qquad \frac{\partial \mathbf{r}}{\partial u^3} \cdot \nabla u^2 = 0;$$

and hence that

$$\frac{\partial \mathbf{r}}{\partial u^1}, \frac{\partial \mathbf{r}}{\partial u^2}, \frac{\partial \mathbf{r}}{\partial u^3} \text{ and } \nabla u^1, \nabla u^2, \nabla u^3$$

are reciprocal systems of vectors equivalent to the unitary and reciprocal unitary systems at P.

2. Show that the gradient of a scalar point function U can be expressed on a general co-ordinate system as follows:

$$\begin{aligned} \nabla U &= \nabla u^1 \frac{\partial U}{\partial u^1} + \nabla u^2 \frac{\partial U}{\partial u^2} + \nabla u^3 \frac{\partial U}{\partial u^3} \\ &= \mathbf{a}^1 \frac{\partial U}{\partial u^1} + \mathbf{a}^2 \frac{\partial U}{\partial u^2} + \mathbf{a}^3 \frac{\partial U}{\partial u^3}; \end{aligned}$$

and hence that:

$$\frac{\partial U}{\partial u^1} = \mathbf{a}_1 \cdot \nabla U, \qquad \frac{\partial U}{\partial u^2} = \mathbf{a}_2 \cdot \nabla U, \qquad \frac{\partial U}{\partial u^3} = \mathbf{a}_3 \cdot \nabla U.$$

3. Give the geometrical significance of the results expressed in exercises (1) and (2).

4. Show that on an orthogonal curvilinear co-ordinate system:

$$\begin{aligned} \nabla g_{11} &= \frac{\mathbf{a}_1}{g_{11}} \frac{\partial g_{11}}{\partial u^1} + \frac{\mathbf{a}_2}{g_{22}} \frac{\partial g_{11}}{\partial u^2} + \frac{\mathbf{a}_3}{g_{33}} \frac{\partial g_{11}}{\partial u^3}, \\ \nabla g_{22} &= \frac{\mathbf{a}_1}{g_{11}} \frac{\partial g_{22}}{\partial u^1} + \frac{\mathbf{a}_2}{g_{22}} \frac{\partial g_{22}}{\partial u^2} + \frac{\mathbf{a}_3}{g_{33}} \frac{\partial g_{22}}{\partial u^3}, \\ \nabla g_{33} &= \frac{\mathbf{a}_1}{g_{11}} \frac{\partial g_{33}}{\partial u^1} + \frac{\mathbf{a}_2}{g_{22}} \frac{\partial g_{33}}{\partial u^2} + \frac{\mathbf{a}_3}{g_{33}} \frac{\partial g_{33}}{\partial u^3}; \end{aligned}$$

$$\nabla u^1 = \frac{\mathbf{a}_1}{g_{11}}, \quad \nabla u^2 = \frac{\mathbf{a}_2}{g_{22}}, \quad \nabla u^3 = \frac{\mathbf{a}_3}{g_{33}}.$$

5. With the aid of the equations given in example (4) show that on a right-handed orthogonal curvilinear co-ordinate system:

$$\operatorname{curl} \mathbf{a}_1 = \frac{1}{\sqrt{g}}\left(\mathbf{a}_2\frac{\partial g_{11}}{\partial u^3} - \mathbf{a}_3\frac{\partial g_{11}}{\partial u^2}\right),$$

$$\operatorname{curl} \mathbf{a}_2 = \frac{1}{\sqrt{g}}\left(\mathbf{a}_3\frac{\partial g_{22}}{\partial u^1} - \mathbf{a}_1\frac{\partial g_{22}}{\partial u^3}\right),$$

$$\operatorname{curl} \mathbf{a}_3 = \frac{1}{\sqrt{g}}\left(\mathbf{a}_1\frac{\partial g_{33}}{\partial u^2} - \mathbf{a}_2\frac{\partial g_{33}}{\partial u^1}\right).$$

6. With the aid of the equations given in example (5) and of the expansion formulas:

$$\nabla(u\mathbf{v}) = \nabla u \cdot \mathbf{v} + u \operatorname{div} \mathbf{v},$$

$$\operatorname{div} (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \operatorname{curl} \mathbf{u} - \mathbf{u} \cdot \operatorname{curl} \mathbf{v},$$

show that on an orthogonal curvilinear co-ordinate system:

$$\operatorname{div} \mathbf{a}_1 = \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{11}}}{\partial u^1} + \frac{1}{\sqrt{g_{22}}}\frac{\partial\sqrt{g_{22}}}{\partial u^1} + \frac{1}{\sqrt{g_{33}}}\frac{\partial\sqrt{g_{33}}}{\partial u^1},$$

$$\operatorname{div} \mathbf{a}_2 = \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{11}}}{\partial u^2} + \frac{1}{\sqrt{g_{22}}}\frac{\partial\sqrt{g_{22}}}{\partial u^2} + \frac{1}{\sqrt{g_{33}}}\frac{\partial\sqrt{g_{33}}}{\partial u^2},$$

$$\operatorname{div} \mathbf{a}_3 = \frac{1}{\sqrt{g_{11}}}\frac{\partial\sqrt{g_{11}}}{\partial u^3} + \frac{1}{\sqrt{g_{22}}}\frac{\partial\sqrt{g_{22}}}{\partial u^3} + \frac{1}{\sqrt{g_{33}}}\frac{\partial\sqrt{g_{33}}}{\partial u^3}.$$

CHAPTER VIII

TRANSFORMATION THEORY

§80

Introductory Remarks

The vectors and vector quantities with which we deal in our 3-dimensional space are, as we know, capable of line-vector representation. In fact, we have so emphasized this purely geometrical property as to lead us, perhaps, to regard it as the most important feature of the vector idea. But in reality it is the algebraic rather than the geometric aspect of a vector which is the more fundamental.

It was formerly thought by some zealous advocates of vector analysis that the use of co-ordinate base-systems and the resolution of vectors into components with the introduction of the corresponding measure-numbers on such systems were actually detrimental to a proper presentation and development of the true vector idea. To them the extensive use of co-ordinate systems in vector analysis was anathema. Now, however, the advantages arising from such use of co-ordinate systems are quite generally recognized.

It is the behavior of the measure-numbers of the components of a vector in passing from one co-ordinate system to another which brings most clearly into evidence the real significance of the vector idea. In fact, it is through the transformation theory of the measure-numbers of vectors that we are led to apprehend their significance from a point of view which permits, and even suggests, generalizations of the idea which are of far-reaching importance, leading, as they do, in a quite natural way to an understanding of the tensor idea.

Fundamental in transformation theory is the notion of invariance of scalar and vector quantities to change of co-ordinate system. In passing from one general co-ordinate system to another, vectors in general are Invariants, but base-vectors are not, since they will change in general both in magnitude and direction with change of co-ordinate system.

A vector which does not change in passing from one co-ordinate system to another is called a Fixed Vector.

From the geometrical or physical nature of a scalar or vector quantity it is often possible to predict that it will be an invariant in any change of co-ordinate system. Predetermination of the invariance of certain quantities is a basic aid in the development of transformation theory.

The magnitude of a fixed vector, the scalar products of two and three fixed vectors, the divergence of a fixed vector, the work done by a force in a displacement, and the energy stored per unit volume in a strained elastic medium, are examples of scalar invariants. In addition to the fixed vector, the vector products of two or three fixed vectors, the gradient of a scalar point function, and the curl of a vector point function are examples of vector invariants.

As the transformation theory for the measure-numbers of the components of fixed vectors and for the base-vectors themselves is developed, it will appear that the equations of transformation can be written down with the aid of very simple rules.

§81

Fundamental Transformation Equations for General Co-ordinate Systems

Let u^1, u^2, u^3 and v^1, v^2, v^3 denote sets of general co-ordinates for any point P, and let the corresponding non-coplanar sets of unitary vectors be denoted respectively by $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ and $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, these vectors varying in general, of course, from point to point. We shall speak of the u and v-co-ordinates as belonging to a U and a V system respectively.

The u and v-co-ordinates are supposed related in some way which is expressed by either of the two sets of equations:

$$(1) \qquad \begin{aligned} v^1 &= v^1(u^1, u^2, u^3), \\ v^2 &= v^2(u^1, u^2, u^3), \\ v^3 &= v^3(u^1, u^2, u^3), \end{aligned} \qquad\qquad \begin{aligned} u^1 &= u^1(v^1, v^2, v^3), \\ u^2 &= u^2(v^1, v^2, v^3), \\ u^3 &= u^3(v^1, v^2, v^3), \end{aligned}$$

in which either the u's or the v's may be chosen as independent variables. For example, these equations in the case of an affine transformation with origin fixed would take the forms:

$$(2) \qquad \begin{aligned} v^1 &= a_1^1 u^1 + a_2^1 u^2 + a_3^1 u^3, \\ v^2 &= a_1^2 u^1 + a_2^2 u^2 + a_3^2 u^3, \\ v^3 &= a_1^3 u^1 + a_2^3 u^2 + a_3^3 u^3, \end{aligned} \qquad\qquad \begin{aligned} u^1 &= b_1^1 v^1 + b_2^1 v^2 + b_3^1 v^3, \\ u^2 &= b_1^2 v^1 + b_2^2 v^2 + b_3^2 v^3, \\ u^3 &= b_1^3 v^1 + b_2^3 v^2 + b_3^3 v^3, \end{aligned}$$

where the a and b-coefficients are constants.

If the co-ordinates of any point Q infinitely close to P be denoted by $u^1 + du^1$, $u^2 + du^2$, $u^3 + du^3$ on the U-system, and by $v^1 + dv^1$, $v^2 + dv^2$, $v^3 + dv^3$ on the V-system, then du^1, du^2, du^3 and dv^1, dv^2, dv^3 can be regarded as the relative co-ordinates of Q with respect to P on the U and V-systems respectively. Let the position-vector of Q with respect to P be denoted by $d\mathfrak{s}$. Then, noting that $d\mathfrak{s}$ is an invariant and using the summation convention, we can write:

$$d\mathfrak{s} = \mathfrak{a}_i du^i \qquad (i = 1,\ 2,\ 3); \tag{3}$$

$$d\mathfrak{s} = \mathfrak{b}_l dv^l, \qquad (l = 1,\ 2,\ 3). \tag{4}$$

Throughout the present chapter it is to be understood in future that all literal indices range over the values 1, 2, 3.

If the reciprocal unitary vectors be denoted as usual by superscripts, we can also write:

$$d\mathfrak{s} = \mathfrak{a}^\lambda du_\lambda, \tag{5}$$

$$d\mathfrak{s} = \mathfrak{b}^\rho dv_\rho, \tag{6}$$

where the du's and dv's are, of course, linear functions of the differentials of the co-ordinates, which are in general non-integrable.

In the present article it is proposed to find the fundamental transformation equations for the passage from the U to the V-system, and from the V to the U-system:

(a) Transformation equations for the differentials of the co-ordinates.

The required equations for this case are obtained directly from equations (1) by differentiation and, making use of the summation convention, can be expressed as follows:

$$dv^l = \frac{\partial v^l}{\partial u^i} du^i, \qquad du^i = \frac{\partial u^i}{\partial v^l} dv^l. \tag{I}$$

The Jacobians $(J,\ K)$ of these transformations are the determinants:

$$J = \begin{vmatrix} \dfrac{\partial v^1}{\partial u^1} & \dfrac{\partial v^1}{\partial u^2} & \dfrac{\partial v^1}{\partial u^3} \\ \dfrac{\partial v^2}{\partial u^1} & \dfrac{\partial v^2}{\partial u^2} & \dfrac{\partial v^2}{\partial u^3} \\ \dfrac{\partial v^3}{\partial u^1} & \dfrac{\partial v^3}{\partial u^2} & \dfrac{\partial v^3}{\partial u^3} \end{vmatrix}; \tag{1a}$$

$$K = \begin{vmatrix} \dfrac{\partial u^1}{\partial v^1} & \dfrac{\partial u^1}{\partial v^2} & \dfrac{\partial u^1}{\partial v^3} \\ \dfrac{\partial u^2}{\partial v^1} & \dfrac{\partial u^2}{\partial v^2} & \dfrac{\partial u^2}{\partial v^3} \\ \dfrac{\partial u^3}{\partial v^1} & \dfrac{\partial u^3}{\partial v^2} & \dfrac{\partial u^3}{\partial v^3} \end{vmatrix}. \tag{2a}$$

Since the u-co-ordinates are supposed mutually independent, and likewise the v-co-ordinates, neither of these determinants can vanish identically. The element $\partial v^l/\partial u^i$ in the l row and i column of J is, by the rules for solving linear homogeneous equations, equal to

the cofactor of the element $\partial u^i/\partial v^l$ in the i row and l column of K divided by K; and vice versa. It follows from equations (I) that:

$$\frac{\partial u^i}{\partial v^l}\frac{\partial v^l}{\partial u^k}du^k = du^i;$$

and hence, by equating coefficients:

$$\frac{\partial u^i}{\partial v^l}\frac{\partial v^l}{\partial u^k} = \delta^i_k, \tag{3a}$$

where $\delta^i_k = 1$, if $k = i$; and $\delta^i_k = 0$, if $k \neq i$. Furthermore, upon forming the product of J and K by the rule for the multiplication of determinants of the same order, and taking account of equation (3a), we find:

$$JK = 1. \tag{4a}$$

The partial derivatives in equation (I) will, of course, vary in general from point to point, but in the case of affine transformations they will be constants. Apart from this difference the equations of transformation for the differentials of general co-ordinates are of the same type as those expressed by equations (2) for the affine co-ordinates themselves.

(b) Transformation equations for the unitary vectors.

From equations (3) and (4) we have:

$$\mathbf{b}_l dv^l = \mathbf{a}_i du^i;$$

therefore, taking account of equations (I):

$$\mathbf{b}_l dv^l \quad \mathbf{a}_i\frac{\partial u^i}{\partial v^l}dv^l, \qquad\qquad \mathbf{a}_i du^i = \mathbf{b}_l\frac{\partial v^l}{\partial u^i}du^i.$$

Since these equations are valid for all values of the dv's and du's, it follows that:

$$\mathbf{b}_l = \frac{\partial u^i}{\partial v^l}\mathbf{a}_i, \qquad\qquad \mathbf{a}_i = \frac{\partial v^l}{\partial u^i}\mathbf{b}_l. \tag{II}$$

Incidentally, by scalar multiplication of the second of these equations by $\mathbf{b}^k$, and of the first by $\mathbf{a}^k$, we find:

$$\frac{\partial v^l}{\partial u^i} = \mathbf{a}_i \cdot \mathbf{b}^l; \tag{1b}$$

$$\frac{\partial u^i}{\partial v^l} = \mathbf{b}_l \cdot \mathbf{a}^i. \tag{2b}$$

(c) Transformation equations for the reciprocal differentials. The du_λ's and dv_ρ's in equations (5) and (6) are called for convenience Reciprocal Differentials. From these equations we have:

$$\mathfrak{b}^\rho dv_\rho = \mathfrak{a}^\lambda du_\lambda.$$

Upon scalar multiplication of this equation by $\mathfrak{b}_k$ and by $\mathfrak{a}_k$ we get:

$$dv_\rho = \mathfrak{b}_\rho \cdot \mathfrak{a}^\lambda du_\lambda, \qquad du_\lambda = \mathfrak{a}_\lambda \cdot \mathfrak{b}^\rho dv_\rho.$$

Hence, taking account of equations (1b) and (2b), we have:

$$\text{(III)} \qquad dv_\rho = \frac{\partial u^\lambda}{\partial v^\rho} du_\lambda, \qquad du_\lambda = \frac{\partial v^\rho}{\partial u^\lambda} dv_\rho.$$

(d) Transformation equations for the reciprocal unitary vectors. From equations (5) and (6) we have:

$$\mathfrak{b}^\rho dv_\rho = \mathfrak{a}^\lambda du_\lambda.$$

This equation, with equations (III), gives:

$$\mathfrak{b}^\rho dv_\rho = \mathfrak{a}^\lambda \frac{\partial v^\rho}{\partial u^\lambda} dv_\rho, \qquad \mathfrak{a}^\lambda du_\lambda = \mathfrak{b}^\rho \frac{\partial u^\lambda}{\partial v^\rho} du_\lambda.$$

Since we can consider all the du's, or all the dv's, as arbitrary, therefore:

$$\text{(IV)} \qquad \mathfrak{b}^\rho = \frac{\partial v^\rho}{\partial u^\lambda} \mathfrak{a}^\lambda, \qquad \mathfrak{a}^\lambda = \frac{\partial u^\lambda}{\partial v^\rho} \mathfrak{b}^\rho.$$

(e) Transformation equations for the coefficients of the fundamental differential quadratic form:

$$\overline{ds}^2 = g_{ij} du^i du^j.$$

Corresponding to this form, introduced in Art. 75, for the U-system, we have for the V-system:

$$\overline{ds}^2 = h_{lm} dv^l dv^m,$$

where h_{lm} corresponds to g_{ik}. It is now desired to find the relationship of these coefficients.

From equations (II) above we get:

$$\mathfrak{b}_l \cdot \mathfrak{b}_m = \frac{\partial u^i}{\partial v^l} \mathfrak{a}_i \cdot \mathfrak{b}_m, \qquad \mathfrak{a}_i \cdot \mathfrak{a}_j = \frac{\partial v^l}{\partial u^i} \mathfrak{b}_l \cdot \mathfrak{a}_j,$$
$$= \frac{\partial u^i}{\partial v^l} \frac{\partial u^j}{\partial v^m} \mathfrak{a}_i \cdot \mathfrak{a}_j, \qquad = \frac{\partial v^l}{\partial u^i} \frac{\partial v^m}{\partial u^j} \mathfrak{b}_l \cdot \mathfrak{b}$$

Hence:

$$\text{(V)} \qquad h_{lm} = \frac{\partial u^i}{\partial v^l} \frac{\partial u^j}{\partial v^m} g_{ij}, \qquad g_{ij} = \frac{\partial v^l}{\partial u^i} \frac{\partial v^m}{\partial u^j} h_{lm}.$$

(f) Transformation equations for the coefficients of the reciprocal differential quadratic form:

$$\overline{ds}^2 = g^{\lambda\mu} du_\lambda du_\mu.$$

Corresponding to this form, introduced in Art. 75 for the U-system, we have for the V-system:

$$\overline{ds}^2 = h^{\rho\sigma} dv_\rho dv_\sigma,$$

where $h^{\rho\sigma}$ corresponds to $g^{\lambda\mu}$.

From equations (IV) above we get:

$$\mathfrak{b}^\rho \cdot \mathfrak{b}^\sigma = \frac{\partial v^\rho}{\partial u^\lambda} \mathfrak{a}^\lambda \cdot \mathfrak{b}^\sigma \qquad \mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu = \frac{\partial u^\lambda}{\partial v^\rho} \mathfrak{b}^\rho \cdot \mathfrak{a}^\mu$$

$$= \frac{\partial v^\rho}{\partial u^\lambda} \frac{\partial v^\sigma}{\partial u^\mu} \mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu, \qquad = \frac{\partial u^\lambda}{\partial v^\rho} \frac{\partial u^\mu}{\partial v^\sigma} \mathfrak{b}^\rho \cdot \mathfrak{b}^\sigma.$$

Hence:

(VI) $$h^{\rho\sigma} = \frac{\partial v^\rho}{\partial u^\lambda} \frac{\partial v^\sigma}{\partial u^\mu} g^{\lambda\mu}, \qquad g^{\lambda\mu} = \frac{\partial u^\lambda}{\partial v^\rho} \frac{\partial u^\mu}{\partial v^\sigma} h^{\rho\sigma}.$$

(g) Transformation equations for the determinant of the coefficients of the fundamental differential quadratic form.

For the U-system the determinant in question is:

(1g) $$\begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix}$$

The corresponding determinant for the V-system will be denoted by h, so that:

(2g) $$\begin{vmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{vmatrix}$$

It is now required to express h in terms of g, and vice versa.

For the element h_{lm} of the l row and the m column of h we have by equation (V):

$$h_{lm} = \frac{\partial u^i}{\partial v^l} \frac{\partial u^j}{\partial v^m} g_{ij}.$$

Now, remembering the definition of the determinant K given above, we have:

$$\frac{\partial u^i}{\partial v^l} g_{ij} = \frac{\partial u^1}{\partial v^l} g_{1j} + \frac{\partial u^2}{\partial v^l} g_{2j} + \frac{\partial u^3}{\partial v^l} g_{3j}$$

= the sum of the products of the elements in the l column of K into the corresponding elements of the j column of g

= the element in the l column and j row of the product of the determinant K by the determinant g.

It follows then that:

$$\frac{\partial u^j}{\partial v^m}\frac{\partial u^i}{\partial v^l}g_{ij} = \frac{\partial u^1}{\partial v^m}\frac{\partial u^i}{\partial v^l}g_{i1} + \frac{\partial u^2}{\partial v^m}\frac{\partial u^i}{\partial v^l}g_{i2} + \frac{\partial u^3}{\partial v^m}\frac{\partial u^i}{\partial v^l}g_{i3}$$

= the sum of the products of the elements in the m column of K into the corresponding element of the l column of Kg

= the element in the l row and m column of the determinant K^2g.

Hence, the determinant h must be equal to the determinant g into the square of the determinant K. In a similar manner the determinant g can be shown to be equal to the determinant h into the square of the determinant J, defined in Art. 81. The required equations of transformation will therefore be:

(VII) $$h = K^2g, \qquad g = J^2h.$$

(h) **Transformation equations for the determinant of the coefficients of the reciprocal differential quadratic form.**

For the U-system the determinant now in question is:

(1h) $$g' = \begin{vmatrix} g^{11} & g^{12} & \\ g^{21} & g^{22} & \\ & g^{32} & \end{vmatrix}$$

The corresponding determinant on the V-system will be denoted by h', so that:

(2h) $$h' \quad \begin{matrix} h^{11} & h^{12} & h^{13} \\ h^{21} & h^{22} & h^{23} \\ h^{31} & h^{32} & \end{matrix}$$

It is now required to express h' in terms of g', and vice versa.

Using equations (VI) and following a procedure entirely analogous to that followed in section (g) above, it is then found that:

(VIII) $$h' = J^2g', \qquad g' = K^2h'$$

§82

Covariant and Contravariant Quantities

Inspection of the equations of transformation (I) to (IV) inclusive of Art. 81 shows that:

The differentials of the co-ordinates and the reciprocal unitary vectors transform in accordance with the same laws, and also that

the reciprocal differentials and the unitary vectors transform in accordance with the same laws; furthermore, that the differentials of the co-ordinates and the reciprocal unitary vectors transform in accordance with laws which are different from those for the transformation of the reciprocal differentials and the unitary vectors.

The differentials of the co-ordinates and the reciprocal unitary vectors are said to transform cogrediently; likewise, the reciprocal differentials and the unitary vectors are said to transform cogrediently; on the other hand, the differentials of the co-ordinates and also the reciprocal unitary vectors are said to transform contragrediently to the reciprocal differentials and to the unitary vectors.

Any three quantities so related as to transform cogrediently with the unitary vectors are called Covariant Quantities.

Any three quantities so related as to transform contragrediently to the unitary vectors are called Contravariant Quantities.

Thus, for example, the reciprocal differentials are covariant, and the reciprocal unitary vectors are contravariant quantities.

By way of further example, if a set of three quantities, when expressed on a U-system of co-ordinates, have values denoted respectively by A^1, A^2, A^3, and, when expressed on a V-system, have corresponding values B^1, B^2, B^3, and if:

$$(1) \qquad B^l = \frac{\partial v^l}{\partial u^i} A^i, \qquad A^i \quad \frac{\partial u^i}{\partial v^l} B^l, \qquad (i, l = 1, 2, 3),$$

then the three quantities will be contravariant quantities; in fact, it can easily be shown that they are contravariant measure-numbers of a vector.

The notation which we have been using has anticipated the need for distinguishing between covariant and contravariant quantities. In this notation covariant quantities have been designated by subscripts and contravariant quantities by superscripts, and, occasionally, where identifying indices have been used with quantities which are neither covariant nor contravariant, the indices have been enclosed in parentheses to indicate this fact. There has been one important exception, however, to this procedure, in that general co-ordinates are always identified by superscripts although, unless the co-ordinates are affine, these co-ordinates are neither covariant nor contravariant quantities.

It is to be specially emphasized that the two members of every equation must balance with respect to their covariant and contravariant properties, or, more conveniently expressed, the covariant

and contravariant dimensions of the two members must be the same, just as the ordinary dimensions of any equation must be the same.

In conformity with our notation in estimating the covariant or contravariant dimensions of any expression, a covariant and a contravariant index which are the same may be considered to cancel each other, provided they both affect quantities which are both above or both below a divisor line, and an index of either sort, covariant or contravariant, affecting a quantity below the divisor line is to be reckoned as equivalent to the same index of the opposite sort above the divisor line.

With this understanding the equations of the present and of the preceding chapter will be found to balance with respect to their covariant and contravariant dimensions. For example, the right-hand member of the first of equations (1) above should be indicated by the notation as contravariant with respect to the index l, since the left-hand member is so indicated, and this is actually the case, since the index i occurs as a contravariant index once above and once below the divisor line, producing cancellation, while the index l occurs once as a contravariant index above the divisor line, so that the right-hand member is, on the whole, indicated as contravariant with respect to the index l.

As a somewhat more complicated example, let us consider the following transformation equation which was found above for the metrical coefficients:

$$h_{lm} = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m} g_{ij}.$$

Here, the notation on the left indicates that the expression on the right is covariant with respect to each of the indices l and m, and inspection shows, in accordance with the notation, that it is both covariant and contravariant once with respect to each of the indices i and j, and that it is covariant once with respect to each of the indices l and m, and hence, on the whole, is indicated as covariant with respect to each of the indices l and m.

Later, in dealing with tensors, we shall have frequent occasion to estimate the covariant and contravariant dimensions of far more complicated expressions with respect to a large number of indices, but the principles which will be involved are the same as those set forth above.

§83

The Affine Transformation Group

The general transformation equations (1), Art. 81, represent what is called by mathematicians a general point transformation group. Since the only restrictions placed upon the u and v-functions are that they shall be analytic, they must be supposed to involve an infinite number of parameters whose variations permit of the equations representing all possible point transformations.

A feature of the general transformation group is that a transformation from u to v-co-ordinates followed by a transformation from the v to w co-ordinates is equivalent to a direct transformation from u to w-co-ordinates. This is the fundamental Group Property.

An important sub-group is the linear substitution group, called the Homogeneous Affine Group, represented by the homogeneous linear transformation equations:

$$
(1) \qquad
\begin{aligned}
v^1 &= a_1^1 u^1 + a_2^1 u^2 + a_3^1 u^3, \\
v^2 & \quad a_1^2 u^1 + a_2^2 u^2 + \\
v^3 &= a_1^3 u^1 + a_2^3 u^2 + a_3^3 u^3;
\end{aligned}
\qquad\qquad
\begin{aligned}
u^1 &= b_1^1 v^1 + b_2^1 v^2 + b_3^1 v^3, \\
u^2 &= b_1^2 v^1 + b_2^2 v^2 + b_3^2 v^3, \\
u^3 &= b_1^3 v^1 + b_2^3 v^2 + b_3^3 v^3
\end{aligned}
$$

This sub-group involves but nine independent parameters, the a-coefficients, or the b-coefficients which can be expressed in terms of them, and by variation of these nine parameters all possible transformations with origin fixed belonging to the affine group are expressible by equations (1).

As was seen in section (b), Art. 72, any affine transformation with origin fixed can be conveniently expressed with the aid of a corresponding dyadic. If $\mathbf{\Phi}$ and $\mathbf{\Psi}$ denote the dyadics through which two affine transformations with origin fixed can be expressed, and if $\mathbf{r}$ denote the position vector of the point whose co-ordinates are to be transformed, then the fundamental group property of affine transformations with origin fixed is expressed by the equation:

$$\mathbf{\Psi} \cdot (\mathbf{\Phi} \cdot \mathbf{r}) = (\mathbf{\Psi} \cdot \mathbf{\Phi}) \cdot \mathbf{r}.$$

All the transformation equations for general co-ordinates, found in Art. 81, are, of course, also valid for affine co-ordinates. Furthermore, on account of the constancy of the coefficients in equations (1) above for the transformation of affine co-ordinates, those for the differentials of the affine co-ordinates, obtained from equations (1) by differentiation, must be of precisely the same form as those for the co-ordinates themselves.

Consequently, from equations (I), Art. 81, the equations of transformation for affine co-ordinates can be expressed in the forms:

$$\text{(I)} \qquad v^l = \frac{\partial v^l}{\partial u^i} u^i, \qquad\qquad u^i = \frac{\partial u^i}{\partial v^l} v^l.$$

§84
Orthogonal Transformations

There are two orthogonal Cartesian transformations which are of special importance. Since on all rectangular Cartesian systems the unitary and reciprocal unitary vectors determine identical systems of axes, the distinction between covariant and contravariant transformations does not exist for them, and they therefore do not require a covariant and contravariant notation. Accordingly, to designate rectangular Cartesian co-ordinates we use x, y, z in place of u^1, u^2, u^3, and x', y', z' in place of v^1, v^2, v^3.

Consider now the transformation equations for rectangular Cartesian co-ordinates:

$$\text{(1)} \qquad \begin{aligned} x' &= a_1^1 x + a_2^1 y + a_3^1 z, \\ y' &= a_1^2 x + a_2^2 y + a_3^2 z, \\ z' &= a_1^3 x + a_2^3 y + a_3^3 z, \end{aligned} \qquad\qquad \begin{aligned} x &= b_1^1 x' + b_2^1 y' + b_3^1 z', \\ y &= b_1^2 x' + b_2^2 y' + b_3^2 z', \\ z &= b_1^3 x' + b_2^3 y' + b_3^3 z', \end{aligned}$$

subject to the condition of orthogonality:

$$\text{(2)} \qquad x'^2 + y'^2 + z'^2 = x^2 + \qquad + z^2.$$

Inserting in equation (2) the expressions for x', y', z' given by equations (1) and comparing coefficients, we find the relations:

$$\text{(3)} \qquad \begin{aligned} (a_1^1)^2 + (a_1^2)^2 + (a_1^3)^2 &= 1, \\ (a_2^1)^2 + (a_2^2)^2 + (a_2^3)^2 &= 1, \\ (a_3^1)^2 + (a_3^2)^2 + (a_3^3)^2 &= 1, \end{aligned} \qquad \begin{aligned} a_2^1 a_3^1 + a_2^2 a_3^2 + a_2^3 a_3^3 &= 0, \\ a_3^1 a_1^1 + a_3^2 a_1^2 + a_3^3 a_1^3 &= 0, \\ a_1^1 a_2^1 + a_1^2 a_2^2 + a_1^3 a_2^3 &= 0; \end{aligned}$$

and there will be six similar relations in which the b's take the places of the a's.

The determinants (J, K) of the transformations are as follows:

$$\begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} \qquad\qquad K \quad \begin{vmatrix} b_1^1 & b_2^1 & b_3^1 \\ b_1^2 & b_2^2 & b_3^2 \\ b_1^3 & b_2^3 & b_3^3 \end{vmatrix}$$

From equations (1), with the aid of the relations (3), we find that:

$$\text{(4)} \qquad b_j^i = a_i^j, \qquad\qquad (i, j = 1, 2, 3).$$

In virtue of the relations (4) it follows that $K = J$, and since by equation (4a), Art. 81, the product of J and K must equal unity, assuming $J \neq 0$, we can write:

$$J^2 = 1,$$

and, therefore:

$$J = \pm 1.$$

It will now be shown that if $J = +1$ for a given set of coefficients in equations (1), then these equations represent a rotation and constitute a sub-group; if $J = -1$, a transformation, called a perversion, consisting of an inversion followed by a rotation.

For the identical transformation:

$$x' = x,\ y' = y,\ z' = z,$$

each of the coefficients a_1^1, a_2^2, a_3^3, will be equal to unity and each of the others will be equal to zero. Hence, for an identical transformation:

$$J = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = +1.$$

But for the inversion transformation:

$$x' = -x,\ y' = -y,\ z' = -x,$$

each of the coefficients a_1^1, a_2^2, a_3^3, will be equal to minus unity and each of the others will be equal to zero, and hence $J = -1$. Now, since any rotation can be obtained from the identical transformation, for which $J = +1$, by a continuous variation of the coefficients, and since J^2 must for any rotation be equal to $+1$, it follows that:

$$J = \begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} = +1,$$

for the rotation group. Again, since any perversion can be obtained from the identical transformation by an inversion, for which $J = -1$, followed by a rotation, and since J^2 must be equal to $+1$, it follows that:

$$J = \begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} = -1,$$

for the perversion transformation.

§85

Transformation Equations for the Measure-Numbers of a Vector and of a Vector Product

Let **p** and **q** respectively denote on the U and the V-systems a fixed vector associated with a field point P. We can express both **p** and **q** in alternative forms as follows:

$$\mathbf{p} = p^i \mathbf{a}_i = p_\lambda \mathbf{a}^\lambda;$$
$$\mathbf{q} = q^l \mathbf{b}_l = q_\lambda \mathbf{b}^\lambda.$$

Since the vector in question is a fixed vector, it will be an invariant, and hence: $\mathbf{p} = \mathbf{q}$. By choosing the infinitesimal vector $d\mathbf{s}$ in the direction of **p**, we can write:

$$\mathbf{p} = \mathbf{q} = k d\mathbf{s},$$

where k is a scalar multiplier, which, of course, is invariant under the transformation. It follows, therefore, that the measure-numbers of **p** and **q** must transform in the same manner as those of $d\mathbf{s}$. Hence, from equations (I) and (III), Art. 81, we have:

$$\text{(I)} \qquad q^l = \frac{\partial v^l}{\partial u^i} p^i, \qquad p^i = \frac{\partial u^i}{\partial v^l} q^l;$$

$$\text{(II)} \qquad q_\rho = \frac{\partial u^\lambda}{\partial v^\rho} p_\lambda, \qquad p_\lambda = \frac{\partial v^\rho}{\partial u^\lambda} q_\rho;$$

these are the required equations of transformation for the measure-numbers of a vector.

Now, let **p** represent the vector product of two vectors **r** and **s**. These two vectors can be expressed in the alternative forms:

$$\text{(1)} \qquad \mathbf{r} = r^1\mathbf{a}_1 + r^2\mathbf{a}_2 + r^3\mathbf{a}_3 = r_1\mathbf{a}^1 + r_2\mathbf{a}^2 + r_3\mathbf{a}^3;$$

$$\text{(2)} \qquad \mathbf{s} = s^1\mathbf{a}_1 + s^2\mathbf{a}_2 + s^3\mathbf{a}_3 = s_1\mathbf{a}^1 + s_2\mathbf{a}^2 + s_3\mathbf{a}^3.$$

With the aid of these equations it is now proposed to find the contravariant and covariant measure-numbers of **p** in terms of those of **r** and **s**.

From equations (1) and (2) we find that:

$$\text{(3)} \qquad \mathbf{p} = \mathbf{r}\times\mathbf{s} = [\mathbf{a}^1\,\mathbf{a}^2\,\mathbf{a}^3]\,(r_2s_3 - r_3s_2)\,\mathbf{a}_1 + \cdots + \cdots;$$

$$\text{(4)} \qquad \mathbf{p} = \mathbf{r}\times\mathbf{s} = [\mathbf{a}_1\,\mathbf{a}_2\,\mathbf{a}_3]\,(r^2s^3 - r^3s^2)\,\mathbf{a}^1 + \cdots + \cdots;$$

in each line the two terms omitted can be obtained from the one given by cyclical permutation of all the indices. From these

equations we infer that the contravariant and the covariant measure-numbers of $\mathbf{p}$ are as follows:

$$
(5) \qquad \begin{aligned}
p^1 &= [\alpha^1\ \alpha^2\ \alpha^3]\,(r_2s_3 - r_3s_2),\\
p^2 &= [\alpha^1\ \alpha^2\ \alpha^3]\,(r_3s_1 - r_1s_3),\\
p^3 &= [\alpha^1\ \alpha^2\ \alpha^3]\,(r_1s_2 - r_2s_1);
\end{aligned}
$$

$$
(6) \qquad \begin{aligned}
p_1 &= [\alpha_1\ \alpha_2\ \alpha_3]\,(r^2s^3 - r^3s^2),\\
p_2 &= [\alpha_1\ \alpha_2\ \alpha_3]\,(r^3s^1 - r^1s^3),\\
p_3 &= [\alpha_1\ \alpha_2\ \alpha_3]\,(r^1s^2 - r^2s^1).
\end{aligned}
$$

With these expressions for p^1, p^2, p^3, and for p_1, p_2, p_3, equations (I) and (II) above will be the appropriate transformation equations for the measure numbers of the vector product $\mathbf{r} \times \mathbf{s}$ provided q^1, q^2, q^3 and q_1, q_2, q_3 represent, respectively, the transformed right-hand members of equations (5) and (6) in order.

It has thus been shown that the measure-numbers of the vector product of any two vectors transform in the same manner as the measure-numbers of an ordinary vector, provided that the measure-numbers of the product are correctly chosen, *i.e.*, in accordance with equations (5) and (6). The notation itself in equations (3) and (4) calls attention to the fact that it is the entire coefficients of α_1, α_2, α_3, and of α^1, α^2, α^3, which are contravariant and covariant respectively, and not the factors which are within parentheses, these being neither contravariant nor covariant.

In the case of a rectangular Cartesian system of axes each of the factors $[\alpha_1\ \alpha_2\ \alpha_3]$ and $[\alpha^1\ \alpha^2\ \alpha^3]$ assumes the value unity and they become identical, but, if an inversion transformation, or a change from right-handed to left-handed axes is contemplated, it must be remembered that these factors will transform into minus unity, and since the measure-numbers of both constituent vectors $\mathbf{r}$ and $\mathbf{s}$ of the vector product $\mathbf{r} \times \mathbf{s}$ will each undergo change of sign, the net change in the measure numbers of $\mathbf{r} \times \mathbf{s}$ in the inversion transformation will consist simply in change of sign, as in the case of an ordinary vector.

Failure to recognize the significance of these factors in an inversion transformation led long ago to the belief that it was necessary to make a distinction between an ordinary vector and a vector which is the vector product of two ordinary vectors, based on the supposed difference in the inversion transformation, whereby the measure-numbers of an ordinary vector were supposed to undergo change of sign while those of the vector product of two ordinary vectors do not. This distinction, really erroneous from the standpoint of

transformation theory, was emphasized by the division of vectors into two classes, viz: Polar Vectors and Axial Vectors. An ordinary vector was called a polar vector, and a vector which is itself the vector product of two ordinary vectors was called an axial vector.

There is, however, a real distinction between so-called polar and axial vectors in the case of a transformation consisting of a change of units of the measure-numbers. Thus, if all three units of the measure-numbers of two ordinary vectors **r** and **s** be changed by a factor, the magnitudes of these vectors will each be changed inversely by the same factor, but the magnitudes of their vector product will be changed inversely by the square of this factor.

Transformation Equations for Base-Vector Dyads

A typical fundamental dyad on the U-system is denoted by $\mathfrak{a}_i\mathfrak{a}_j$; the indices i and j can assume, of course, any of the integer values 1, 2, 3; on a V-system of co-ordinates let the typical fundamental dyad be denoted by $\mathbf{b}_l\mathbf{b}_m$. It is required to find the transformation equations for the corresponding systems of dyads.

With the aid of equations (II), Art. 81, it is found directly that:

$$\text{(I)} \qquad \mathbf{b}_l\mathbf{b}_m = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m}\mathfrak{a}_i\mathfrak{a}_j, \qquad \mathfrak{a}_i\mathfrak{a}_j = \frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j}\mathbf{b}_l\mathbf{b}_m.$$

These transformation equations show that the $\mathfrak{a}_i\mathfrak{a}_j$-system is doubly covariant.

In like manner we find, with the aid of equation (IV), Art. 81, for the system of dyads whose antecedents and consequents are reciprocal unitary vectors:

$$\text{(II)} \qquad \mathbf{b}^\rho\mathbf{b}^\sigma = \frac{\partial v^\rho}{\partial u^\lambda}\frac{\partial v^\sigma}{\partial u^\mu}\mathfrak{a}^\lambda\mathfrak{a}^\mu, \qquad \mathfrak{a}^\lambda\mathfrak{a}^\mu = \frac{\partial u^\lambda}{\partial v^\rho}\frac{\partial u^\mu}{\partial v^\sigma}\mathbf{b}^\rho\mathbf{b}^\sigma.$$

These transformation equations show that the $\mathfrak{a}^\lambda\mathfrak{a}^\mu$-system is doubly contravariant.

For mixed fundamental dyads, with the aid of equations (II) and (IV), Art. 81, we find.

$$\text{(III)} \qquad \mathbf{b}_l\mathbf{b}^\sigma = \frac{\partial u^i}{\partial v^l}\frac{\partial v^\sigma}{\partial u^\mu}\mathfrak{a}_i\mathfrak{a}^\mu, \qquad \mathfrak{a}_i\mathfrak{a}^\mu = \frac{\partial v^l}{\partial u^i}\frac{\partial u^\mu}{\partial v^\sigma}\mathbf{b}_l\mathbf{b}^\sigma;$$

$$\text{(IV)} \qquad \mathbf{b}^\rho\mathbf{b}_m = \frac{\partial v^\rho}{\partial u^\lambda}\frac{\partial u^j}{\partial v^m}\mathfrak{a}^\lambda\mathfrak{a}_j, \qquad \mathfrak{a}^\lambda\mathfrak{a}_j = \frac{\partial u^\lambda}{\partial v^\rho}\frac{\partial v^m}{\partial u^j}\mathbf{b}^\rho\mathbf{b}_m.$$

The dyad $\mathfrak{a}_i\mathfrak{a}^\mu$ is covariant with respect to the index i and contravariant with respect to the index μ, and the dyad $\mathfrak{a}^\lambda\mathfrak{a}_j$ is contravariant with respect to the index λ and covariant with respect to the index j.

§87

Transformation Equations for Various Forms of a Dyadic

In Art. 74 it was shown that a dyadic $\mathbf{\Phi}$ can be expressed in the following forms:

$$\begin{aligned}\mathbf{\Phi} &= a^{ij}\mathfrak{a}_i\mathfrak{a}_j\\ &= a_{\lambda}^{\cdot j}\mathfrak{a}^\lambda\mathfrak{a}_j\\ &= a^{i}_{\cdot\mu}\mathfrak{a}_i\mathfrak{a}^\mu\\ &= a_{\lambda\mu}\mathfrak{a}^\lambda\mathfrak{a}^\mu,\end{aligned}$$

on a U-system of co-ordinates. Let the corresponding forms on a V-system of co-ordinates be expressed as follows:

$$\begin{aligned}\mathbf{\Phi} &= b^{lm}\mathfrak{b}_l\mathfrak{b}_m\\ &= b_{\rho}^{\cdot m}\mathfrak{b}^\rho\mathfrak{b}_m\\ &= b^{l}_{\cdot\sigma}\mathfrak{b}_l\mathfrak{b}^\sigma\\ &= b_{\rho\sigma}\mathfrak{b}^\rho\mathfrak{b}^\sigma.\end{aligned}$$

The transformation equations for the a and b-coefficients in the above forms are required.

Since $\mathbf{\Phi}$ is an invariant, we can write:

$$b^{lm}\mathfrak{b}_l\mathfrak{b}_m = a^{ij}\mathfrak{a}_i\mathfrak{a}_j.$$

From this equation, with the aid of equations (II), Art. 81, we find:

$$b^{lm}\mathfrak{b}_l\mathfrak{b}_m = a^{ij}\frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j}\mathfrak{b}_l\mathfrak{b}_m,$$

$$a^{ij}\mathfrak{a}_i\mathfrak{a}_j = b^{lm}\frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m}\mathfrak{a}_i\mathfrak{a}_j.$$

By the criteria for the equality of two dyadics it then follows that:

$$\text{(I)}\qquad b^{lm} = \frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j}a^{ij},\qquad a^{ij} = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m}b^{lm}.$$

In like manner it can be shown that the mixed and the covariant coefficients of $\mathbf{\Phi}$ transform as follows:

$$\text{(II)}\qquad b_{\rho}^{\cdot m} = \frac{\partial u^\lambda}{\partial v^\rho}\frac{\partial v^m}{\partial u^j}a_{\lambda}^{\cdot j},\qquad a_{\lambda}^{\cdot j} = \frac{\partial v^\rho}{\partial u^\lambda}\frac{\partial u^j}{\partial v^m}b_{\rho}^{\cdot m};$$

$$\text{(III)}\qquad b^{l}_{\cdot\sigma} = \frac{\partial v^l}{\partial u^i}\frac{\partial u^\mu}{\partial v^\sigma}a^{i}_{\cdot\mu},\qquad a^{i}_{\cdot\mu} = \frac{\partial u^i}{\partial v^l}\frac{\partial v^\sigma}{\partial u^\mu}b^{l}_{\cdot\sigma};$$

$$\text{(IV)}\qquad b_{\rho\sigma} = \frac{\partial u^\lambda}{\partial v^\rho}\frac{\partial u^\mu}{\partial v^\sigma}a_{\lambda\mu},\qquad a_{\lambda\mu} = \frac{\partial v^\rho}{\partial u^\lambda}\frac{\partial v^\sigma}{\partial u^\mu}b_{\rho\sigma}.$$

§88

Differential Invariants on General Co-ordinate Systems

The differential invariants which will be considered in the present article are as follows:

The gradient of a scalar point function.
The operator ∇.
The divergence of a vector point function.
The Lame′ operator Δ.
The curl of a vector point function.

(a) **The gradient of a scalar point function.**

The gradient of a scalar point function U is a fixed vector and therefore an invariant. On a rectangular Cartesian system of co-ordinates it can be expressed in the form:

$$\nabla U = \mathbf{i}\frac{\partial U}{\partial x} + \mathbf{j}\frac{\partial U}{\partial y} + \mathbf{k}\frac{\partial U}{\partial z}.$$

By transformation it is quite possible to obtain a corresponding form for any general system of co-ordinates, but it is simpler at present to proceed as follows:

Let $P\ (u^1, u^2, u^3)$ be a field point and $d\mathbf{r}$ the position-vector with respect to P of an infinitely near point $Q(u^1 + du^1,\ u^2 + du^2,\ u^3 + du^3)$. Then the increase in U, denoted by dU, in passing from P to Q will be expressed by the equation:

$$dU = \nabla U \cdot d\mathbf{r}.$$

But we have also:

$$dU = \frac{\partial U}{\partial u^\lambda}\, du^\lambda.$$

Therefore:

$$\nabla U \cdot d\mathbf{r} = \frac{\partial U}{\partial u^\lambda}\, du^\lambda.$$

Expressing $d\mathbf{r}$ on an $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$-base-system at P. we have:

$$d\mathbf{r} = \mathfrak{a}_\lambda du^\lambda,$$

where:

$$du^\lambda = \mathfrak{a}^\lambda \cdot d\mathbf{r}.$$

Hence:

$$\nabla U \cdot d\mathbf{r} = \left(\mathfrak{a}^\lambda \frac{\partial U}{\partial u^\lambda}\right) \cdot d\mathbf{r}.$$

Since this equation is valid for all possible values of $d\mathbf{r}$, it follows that:

$$\nabla U = \mathfrak{a}^{\lambda}\frac{\partial U}{\partial u^{\lambda}} \tag{1a}$$

Here, the gradient of U is expressed on the reciprocal unitary base-system at P; the partial derivatives are its covariant measure-numbers.

With the aid of the first of equations (4), Art. 74, the gradient of U can be expressed on the unitary base-system as follows:

$$\nabla U = \mathfrak{a}_i\frac{\partial U}{\partial u_i}, \tag{2a}$$

where, for its contravariant measure-numbers, we have:

$$\frac{\partial U}{\partial u_i} = g^{\lambda i}\frac{\partial U}{\partial u^{\lambda}}. \tag{3a}$$

By direct transformation of either of the forms (1a) or (2a) it can easily be shown that ∇U is an invariant, but, since ∇U is a fixed vector, this fact is already known.

(b) The operator ∇.

From equations (1a) and (2a) it appears that the operator ∇, which is expressed on an x, y, z-co-ordinate system in the form:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z},$$

must transform so as to give on a general U-co-ordinate system the alternative forms:

$$\nabla = \mathfrak{a}^{\lambda}\frac{\partial}{\partial u^{\lambda}}, \tag{1b}$$

$$\nabla = \mathfrak{a}_i\frac{\partial}{\partial u_i}, \tag{2b}$$

where:

$$\frac{\partial}{\partial u^{\lambda}} = g_{i\lambda}\frac{\partial}{\partial u_i}, \tag{3b}$$

$$\frac{\partial}{\partial u_i} = g^{\lambda i}\frac{\partial}{\partial u^{\lambda}} \tag{4b}$$

Since the gradient of any scalar point function U as well as U itself is an invariant, it follows from equation (1a) or (2a) that the operator ∇ must also be an invariant, as the notation indicates. In passing from a general U to a general V-co-ordinate system, the forms (1b) and (2b) must therefore be invariant, so that:

$$\text{(I)} \qquad \nabla = \mathfrak{a}^{\lambda}\frac{\partial}{\partial u^{\lambda}} = \mathfrak{b}^{\rho}\frac{\partial}{\partial v^{\rho}};$$

$$\text{(II)} \qquad \nabla = \mathfrak{a}_{i}\frac{\partial}{\partial u_{i}} = \mathfrak{b}_{l}\frac{\partial}{\partial v_{l}}.$$

(c) The divergence of a vector point function.

The definition of the divergence of a vector point function given in Art. 36 requires it to be an invariant. If **A** denote such a function, and X, Y, Z its measure-numbers on a rectangular Cartesian system of co-ordinates, then its divergence can, as we know, be expressed in the form:

$$\operatorname{div}\mathbf{A} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}.$$

It is, of course, possible by direct transformation to obtain the corresponding form on a general system of co-ordinates, but it is easier and more instructive to use a more or less geometric method.

By formula (10), Art. 36, the divergence of a vector point function **A** at any point $P(u^1, u^2, u^3)$ is equal to the surface integral of the normal component of **A** over the surface ω bounding an element of volume of any shape divided by the magnitude $d\tau$ of the element. Referring to Fig. 36, p. 166, we take as the volume element one which is delimited by the level surfaces of the co-ordinates passing through the point $P(u^1, u^2, u^3)$ and those passing through the neighboring point $Q(u^1 + du^1, u^2 + du^2, u^3 + du^3)$ where du^1, du^2, du^3 are all positive. As usual, let $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$ denote the unitary vectors at P.

The contribution to the divergence of **A** of the pair of faces of the element for which the u^1-co-ordinates are respectively u^1 and $u^1 + du^1$ is given by the expression:

$$\pm\frac{1}{d\tau}\left[(\mathbf{A}\cdot\mathfrak{a}_2\times\mathfrak{a}_3 du^2du^3)_{u^1+du^1} - (\mathbf{A}\cdot\mathfrak{a}_2\times\mathfrak{a}_3du^2du^3)_{u^1}\right],$$

where the notation signifies that the quantities in parentheses are, respectively, to be evaluated for values of the co-ordinates $u^1 + du^1$, u^2, u^3 and u^1, u^2, u^3, and the $+$ or $-$ sign is to be used according as the angle between the vectors $\mathfrak{a}_2 \times \mathfrak{a}_3$ and $\mathfrak{a}_1$ is acute or obtuse. To quantities of the second order of smallness in du^1 this expression is equal to the following one:

$$\pm\frac{1}{d\tau}\,\frac{\partial}{\partial u^1}(\mathbf{A}\cdot\mathfrak{a}_2\times\mathfrak{a}_3du^1du^2du^3),$$

or:

$$\pm\frac{1}{d\tau}\,\frac{\partial}{\partial u^1}\left[\mathbf{A}\cdot\mathfrak{a}^1(\mathfrak{a}_1\cdot\mathfrak{a}_2\times\mathfrak{a}_3)\,du^1du^2du^3\right],$$

with the aid of equations (1), Art. 19; but, as shown in Art. 77:

$$\pm \mathfrak{a}_1 \cdot \mathfrak{a}_2\mathfrak{a}_3 du^1 du^2 du^3 = d\tau = \sqrt{g}\, du^1 du^2 du^3,$$

where the $+$ or $-$ sign is to be taken according as the angle between the vectors $\mathfrak{a}_2 \times \mathfrak{a}_3$ and $\mathfrak{a}_1$ is acute or obtuse; hence, the contribution to the divergence of **A** of the pair of faces in question is:

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial u^1} (\sqrt{g}\, \mathbf{A} \cdot \mathfrak{a}^1);$$

and since the other two pairs of surfaces of the volume element make similar contributions, we therefore have:

$$\text{(1c)} \qquad \operatorname{div} \mathbf{A} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial u^\lambda} (\sqrt{g}\mathbf{A} \cdot \mathfrak{a}^\lambda),$$

or:

$$\text{(2c)} \qquad \operatorname{div} \mathbf{A} \quad \frac{\ }{\sqrt{g}} \frac{\ }{\partial u^\lambda} (\sqrt{g} A^\lambda),$$

where summation is implied with respect to the index λ, and where A^λ denotes the typical contravariant component of **A**.

(d) The Lame' operator Δ.

If in equation (1c) the vector **A** is the gradient of a scalar point function U, then:

$$\mathbf{A} = \nabla U = \mathfrak{a}^\mu \frac{\partial U}{\partial u^\mu},$$

and, therefore:

$$\nabla \cdot \nabla U = \frac{1}{\sqrt{g}} \frac{\partial}{\partial u^\lambda} \left(\sqrt{g}\, \mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu \frac{\partial U}{\partial u^\mu} \right).$$

Hence, upon writing ΔU for $\nabla \cdot \nabla U$ and $g^{\lambda\mu}$ for $\mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu$, we have:

$$\text{(1d)} \qquad \Delta U = \frac{1}{\sqrt{g}} \frac{\partial}{\partial u^\lambda} \left(\sqrt{g}\, g^{\lambda\mu} \frac{\partial U}{\partial u^\mu} \right).$$

Since ΔU represents the divergence of the gradient of U, which is a fixed vector, we infer at once that ΔU is an invariant. It follows that the Lame' operator[1]:

$$\text{(2d)} \qquad \Delta = \frac{1}{\sqrt{g}} \frac{\partial}{\partial u^\lambda} \left(\sqrt{g}\, g^{\lambda\mu} \frac{\partial}{\partial u^\mu} \right),$$

must also be an invariant.

(e) The curl of a vector point function.

By its definition given in Art. 36 the curl of any point function **A** must be an invariant. If X, Y, Z denote the measure-numbers

[1] In rectangular Cartesian co-ordinates this operator is identical with ∇^2, the Laplacian operator.

of **A** on a rectangular Cartesian system of co-ordinates, then, as we know, the curl of **F** can be expressed in the form:

$$\begin{aligned}\text{curl } \mathbf{A} = &\left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}\right)\mathbf{i}\\ &+ \left(\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}\right)\mathbf{j}\\ &+ \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right)\mathbf{k}.\end{aligned}$$

By direct transformation of this expression it is possible to obtain a corresponding form on a general system of co-ordinates; but the general form can be obtained more easily as follows:

By formula (11), Art. 36, the curl of **A** at any point $P(u^1, u^2, u^3)$ is expressed by the equation:

$$\text{curl } \mathbf{A} = \frac{1}{d\tau}\int \boldsymbol{\varepsilon} \times \mathbf{A}\, d\sigma,$$

where ω denotes an infinitely small closed surface enclosing a differential element of volume of magnitude $d\tau$ within which P is located, $d\sigma$ the magnitude of the area of a differential element of ω, and $\boldsymbol{\varepsilon}$ a unit vector in the direction of an outward drawn normal to ω. The $\mathfrak{a}_1$-component of curl **A** on an $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$-base-system at P can be expressed as follows:

$$\begin{aligned}(\mathfrak{a}^1 \cdot \text{curl } \mathbf{A})\, \mathfrak{a}_1 &= \frac{\mathfrak{a}_1}{d\tau}\int_\omega \mathfrak{a}^1 \cdot \boldsymbol{\varepsilon} \times \mathbf{A}\, d\sigma\\ &= \frac{\mathfrak{a}_1}{d\tau}\int_\omega \boldsymbol{\varepsilon} \cdot \mathbf{A} \times \mathfrak{a}^1 d\sigma\\ &= \mathfrak{a}_1 \text{ div } \mathbf{A} \times \mathfrak{a}^1,\end{aligned}$$

with the aid of equation (10), Art. 36. Now, as was shown in Art. 77:

$$\mathfrak{a}^1 = \frac{\mathfrak{a}_2 \times \mathfrak{a}_3}{[\mathfrak{a}_1\, \mathfrak{a}_2\, \mathfrak{a}_3]} = \pm\frac{\mathfrak{a}_2 \times \mathfrak{a}_3}{\sqrt{g}};$$

and, if the $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$-base-system is right-handed, as we shall assume, the plus sign must be used. Hence:

$$\begin{aligned}(\mathfrak{a}^1 \cdot \text{curl } \mathbf{A})\, \mathfrak{a}_1 &= \mathfrak{a}_1 \text{ div } \frac{\mathbf{A} \times (\mathfrak{a}_2 \times \mathfrak{a}_3)}{\sqrt{g}}\\ &= \mathfrak{a}_1 \text{ div } \frac{1}{\sqrt{g}}(\mathbf{A} \cdot \mathfrak{a}_3\mathfrak{a}_2 - \mathbf{A} \cdot \mathfrak{a}_2\mathfrak{a}_3)\\ &= \frac{\mathfrak{a}_1}{\sqrt{g}}\frac{\partial}{\partial u^\lambda}[(\mathbf{A} \cdot \mathfrak{a}_3\mathfrak{a}_2 - \mathbf{A} \cdot \mathfrak{a}_2\mathfrak{a}_3) \cdot \mathfrak{a}^\lambda],\end{aligned}$$

with the aid of equation (1c) above. Hence, for the final expression for the $\mathbf{a}_1$-component of the curl of $\mathbf{A}$ we shall have:

$$\frac{\mathbf{a}_1}{\sqrt{g}}\left(\frac{\partial \mathbf{A}\cdot\mathbf{a}_3}{\partial u^2} - \frac{\partial \mathbf{A}\cdot\mathbf{a}_2}{\partial u^3}\right).$$

The $\mathbf{a}_2$ and $\mathbf{a}_3$-components can be obtained from the $\mathbf{a}_1$-component by cyclical permutation of the indices 1, 2, 3. Upon adding these components we get as the final result:

$$\begin{aligned} \operatorname{curl}\mathbf{A} = \frac{1}{\sqrt{g}}\Big[&\left(\frac{\partial \mathbf{A}\cdot\mathbf{a}_3}{\partial u^2} - \frac{\partial \mathbf{A}\cdot\mathbf{a}_2}{\partial u^3}\right)\mathbf{a}_1 \\ &+ \left(\frac{\partial \mathbf{A}\cdot\mathbf{a}_1}{\partial u^3} - \frac{\partial \mathbf{A}\cdot\mathbf{a}_3}{\partial u^1}\right)\mathbf{a}_2 \\ &+ \left(\frac{\partial \mathbf{A}\cdot\mathbf{a}_2}{\partial u^1} - \frac{\partial \mathbf{A}\cdot\mathbf{a}_1}{\partial u^2}\right)\mathbf{a}_3\Big], \end{aligned} \tag{1e}$$

or:

$$\begin{aligned} \operatorname{curl}\mathbf{A} = \frac{1}{\sqrt{g}}\Big[&\left(\frac{\partial A_3}{\partial u^2} - \frac{\partial A_2}{\partial u^1}\right)\mathbf{a}_1 \\ &+ \left(\frac{\partial A_1}{\partial u^3} - \frac{\partial A_3}{\partial u^1}\right)\mathbf{a}_2 \\ &+ \left(\frac{\partial A_2}{\partial u^1} - \frac{\partial A_1}{\partial u^2}\right) \quad \Big], \end{aligned} \tag{2e}$$

where the quantities enclosed in parentheses multiplied by $1/\sqrt{g}$ are the contravariant measure-numbers of curl $\mathbf{A}$, and the scalar A's are the covariant measure-numbers of the vector $\mathbf{A}$ itself.

Differential Invariants on Orthogonal Co-ordinate Systems

As a reminder that we are dealing in the present article with orthogonal co-ordinate systems, all identifying indices will be enclosed in parentheses.

The conditions of orthogonality are expressed through the unitary vectors as follows:

$$\mathbf{a}_{(2)}\cdot\mathbf{a}_{(3)} = \mathbf{a}_{(3)}\cdot\mathbf{a}_{(1)} = \mathbf{a}_{(1)}\cdot\mathbf{a}_{(2)} = 0.$$

By the second of equations (9), Art. 73, these conditions require that the reciprocal unitary vectors shall be related to the unitary vectors as follows:

$$\mathbf{a}^{(1)} = \frac{1}{a^2_{(1)}}\mathbf{a}_{(1)}; \qquad \mathbf{a}^{(2)} = \frac{1}{a^2_{(2)}}\mathbf{a}_{(2)}; \qquad \mathbf{a}^{(3)} = \frac{1}{a^2_{(3)}}\mathbf{a}_{(3)}.$$

By their definitions given in Art. 74 the g-coefficients of the differential quadratic forms must therefore have on orthogonal systems the following values:

$$g_{(11)} \quad a_{(1)}^2, \qquad g^{(11)} \quad \frac{}{a_{(1)}^2}, \qquad g_{(23)} = g_{(32)} = g^{(23)} = g^{(32)} = 0,$$

$$g_{(22)} \quad a_{(2)}^2, \qquad g^{(22)} = \frac{1}{a_{(2)}^2}, \qquad g_{(31)} = g_{(13)} = g^{(31)} = g^{(13)} = 0,$$

$$g_{(33)} = a_{(3)}^2; \qquad g^{(33)} = \frac{1}{a_{(3)}^2}; \qquad g_{(12)} = g_{(21)} = g^{(12)} = g^{(21)} = 0.$$

The determinants of these coefficients then assume the forms:

$$\begin{vmatrix} g_{(11)} & 0 & 0 \\ 0 & g_{(22)} & 0 \\ 0 & 0 & g_{(33)} \end{vmatrix} = g_{(11)}g_{(22)}g_{(33)}; \qquad g' = \begin{vmatrix} g^{(11)} & 0 & 0 \\ 0 & g^{(22)} & 0 \\ 0 & 0 & g^{(33)} \end{vmatrix} = g^{(11)}g^{(22)}g^{(33)}.$$

Taking into account these relations, the differential invariant forms found for general co-ordinate systems in Art. 88 are easily seen to reduce to the following forms for orthogonal systems:

(1) $$\nabla U \quad \frac{\mathfrak{a}_{(1)}}{g_{(11)}} \frac{\partial U}{\partial u^{(1)}} + \frac{\mathfrak{a}_{(2)}}{g_{(22)}} \frac{\partial U}{\partial u^{(2)}} + \frac{\mathfrak{a}_{(3)}}{g_{(33)}} \frac{\partial U}{\partial u^{(3)}};$$

(2) $$\nabla = \frac{\mathfrak{a}_{(1)}}{g_{(11)}} \frac{\partial}{\partial u^{(1)}} + \frac{\mathfrak{a}_{(2)}}{g_{(22)}} \frac{\partial}{\partial u^{(2)}} + \frac{\mathfrak{a}_{(3)}}{g_{(33)}} \frac{}{\partial u^{(3)}},$$

(3) $$\begin{aligned} \operatorname{div} \mathbf{A} \quad \frac{1}{\sqrt{g_{(11)}g_{(22)}g_{(33)}}} \Bigg[& \frac{\partial}{\partial u^{(1)}} \left(\sqrt{\frac{g_{(22)}g_{(33)}}{g_{(11)}}}\, \mathbf{A} \cdot \mathfrak{a}_{(1)} \right) \\ & + \frac{\partial}{\partial u^{(2)}} \left(\sqrt{\frac{g_{(33)}g_{(11)}}{g_{(22)}}}\, \mathbf{A} \quad \mathfrak{a}_{(2)} \right. \\ & + \frac{\partial}{\partial u^{(3)}} \left(\sqrt{\frac{g_{(11)}g_{(22)}}{g_{(33)}}}\, \mathbf{A} \cdot \mathfrak{a}_{(3)} \right) \Bigg]; \end{aligned}$$

(4) $$\begin{aligned} \frac{}{\sqrt{g_{(11)}g_{(22)}g_{(33)}}} \Bigg[& \frac{\partial}{\partial u^{(1)}} \left(\sqrt{\frac{g_{(22)}g_{(33)}}{g_{(11)}}} \frac{\partial}{\partial u^{(1)}} \right) \\ & + \frac{\partial}{\partial u^{(2)}} \left(\sqrt{\frac{g_{(33)}g_{(11)}}{g_{(22)}}} \frac{\partial}{\partial u^{(2)}} \right) \\ & \frac{}{\partial u^{(3)}} \left(\sqrt{\frac{g_{(11)}g_{(22)}}{g_{(33)}}} \frac{\partial}{\partial u^{(3)}} \right) \Bigg]; \end{aligned}$$

$$\text{curl A} = \frac{1}{\sqrt{g_{(11)}g_{(22)}g_{(33)}}}\left[\left(\frac{\partial \mathrm{A}\cdot\mathfrak{a}_{(3)}}{\partial u^{(2)}} - \frac{\partial \mathrm{A}\cdot\mathfrak{a}_{(2)}}{\partial u^{(3)}}\right)\mathfrak{a}_{(1)}\right.$$

$$(5) \qquad + \left(\frac{\partial \mathrm{A}\cdot\mathfrak{a}_{(1)}}{\partial u^{(3)}} - \frac{\partial \mathrm{A}\cdot\mathfrak{a}_{(3)}}{\partial u^{(1)}}\right)\mathfrak{a}_{(2)}$$

$$\left. + \left(\frac{\partial \mathrm{A}\cdot\mathfrak{a}_{(2)}}{\partial u^{(1)}} - \frac{\partial \mathrm{A}\cdot\mathfrak{a}_{(1)}}{\partial u^{(2)}}\right)\mathfrak{a}_{(3)}\right].$$

From these forms, which are applicable for any orthogonal system, the corresponding special forms applicable to special orthogonal systems can be written down at once when the corresponding g-coefficients have been determined.

The values of the g-coefficients in these formulas for cylindrical, spherical, and ellipsoidal co-ordinate systems, found in Art. 78, are given in tabular form below:

	Cylindrical Co-ordinates	Spherical Co-ordinates	Ellipsoidal Co-ordinates
	ρ, ϕ, z	r, θ, ϕ	λ, μ, ν
$g_{(11)}$	1	1	$\frac{1}{4}\frac{(\lambda-\mu)(\lambda-\nu)}{(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)}$
$g_{(22)}$			$\frac{1}{4}\frac{(\mu-\nu)(\mu-\lambda)}{(a^2+\mu)(b^2+\mu)(c^2+\mu)}$
$g_{(33)}$		$r^2\sin^2\theta$	$\frac{1}{4}\frac{(\nu-\lambda)(\nu-\mu)}{(a^2+\nu)(b^2+\nu)(c^2+\nu)}$.

EXERCISES ON CHAPTER VIII

1. Show that the transformation equations for the measure-numbers of a vector possess the fundamental group property.

2. Show that the derivatives with respect to the co-ordinates of a scalar point function transform covariantly.

3. Show that the transformation equations for the measure-numbers of a dyadic possess the fundamental group property.

4. Verify the statement: the measure numbers of a dyadic transform in the same manner as products of the measure-numbers of two vectors.

5. An infinitesimal affine transformation is expressed by the equations:

$$\begin{aligned} v^1 &= \epsilon\alpha + (1+\epsilon a_1^1)\,u^1 + \epsilon a_2^1 u^2 + \epsilon a_3^1 u^3, \\ v^2 &= \epsilon\beta + \epsilon a_1^2 u^1 + (1+\epsilon a_2^2)\,u^2 + \epsilon a_3^2 u^3, \\ v^3 &= \epsilon\gamma + \epsilon a_1^3 u^1 + \epsilon a_2^3 u^2 + (1+\epsilon a_3^3)\,u^3, \end{aligned}$$

where ϵ is an infinitesimal and α, β, γ, and the a's are finite constants. Show that the transformation can be decomposed into: a translation + an expansion + a rotation, and that the order in which these component

transformations are taken is immaterial; also show that the Jacobean of the transformation is as follows:

$$\frac{\partial(v^1, v^2, v^3)}{\partial(u^1, u^2, u^3)} = 1 + \epsilon(a_1^1 + a_2^2 + a_3^3).$$

6. In passing from an **i**, **j**, **k**-system to an affine system of axes determined by the constant unitary vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, prove by direct transformation that:

$$\nabla U = \mathbf{i}\frac{\partial U}{\partial x} + \mathbf{j}\frac{\partial U}{\partial y} + \mathbf{k}\frac{\partial U}{\partial z} \text{ becomes } \nabla U = \mathbf{a}^1\frac{\partial U}{\partial u^1} + \mathbf{a}^2\frac{\partial U}{\partial u^2} + \mathbf{a}^3\frac{\partial U}{\partial u^3};$$

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \text{ becomes } \nabla^2 U = g^{\lambda\mu}\frac{\partial^2 U}{\partial u^\lambda \partial u^\mu},$$

where in the last equation summation with respect to the indices λ and μ is implied.

7. Show that the transformation equations for the contravariant and the covariant measure-numbers of an anti-symmetric dyadic can be written in the forms:

$$b^{lm} = \frac{\partial(v^l, v^m)}{\partial(u^i, u^j)}\, a^{ij}, \qquad a^{ij} = \frac{\partial(u^i, u^j)}{\partial(v^l, v^m)}\, b^{lm};$$

$$b_{\rho\sigma} = \frac{\partial(u^\lambda, u^\mu)}{\partial(v^\rho, v^\sigma)}\, a_{\lambda\mu}, \qquad a_{\lambda\mu} = \frac{\partial(v^\rho, v^\sigma)}{\partial(u^\lambda, u^\mu)}\, b_{\rho\sigma};$$

where summation in which no Jacobean is repeated is implied with respect to two indices in each expression.

8. Show that whatever system of co-ordinates may happen to be in use in Euclidean 3-space it is possible by a suitable transformation to find a system for which the metrical coefficients are all constants. Show that this is not true in the case of a non-Euclidean space represented by a non-developable surface.

CHAPTER IX

NON-EUCLIDEAN MANIFOLDS

§90

Fundamental Concepts and Definitions

For the purposes of what may be called pre-relativity physics a vector analysis dealing with vectors of our 3-dimensional Euclidean space was fairly adequate. But modern advances in physics, exemplified particularly by relativity theory, demanded an extension of vector analysis so as to include within its scope a theory of vectors in non-Euclidean space of four dimensions, akin to that which was developed in Art. (31) for the particular case of a surface for which the Gaussian curvature does not vanish.

Fortunately, the necessary pioneer work for such an extension of the scope of vector analysis had already been done by the German mathematicians Karl Friedrich Gauss (1777–1855), Bernhard Riemann (1826–1866), E. Christoffel (1829–1900), and the Italian mathematician G. Ricci (1853–1925). Upon the foundations laid by these investigators the extensions of vector analysis demanded by modern physics have been made by various writers, particularly by the Italian mathematician Levi-Civita (1873–   ), a former pupil of Ricci.

It is the purpose of the present chapter to deal with some of the more important metrical properties of non-Euclidean manifolds having an arbitrary number n of dimensions, to define what is to be understood by a vector in such a manifold, and to indicate how an appropriate vector calculus for such generalized vectors can be developed.

The notation which will be used corresponds in detail with that used in several previous chapters, and is, as previously, specifically designed to bring into evidence covariant and contravariant distinctions.

Throughout the chapter the summation convention is to be understood as operative, and, unless otherwise stated, all summations indicated by the convention are to be considered taken over the integer range $1, 2, \ldots n$.

If to each of a set of n independent variables $u^1, u^2, \ldots, u^n$ a definite value be assigned, then the set of values so obtained is said to determine a point in an n-Dimensional Manifold, and the manifold itself is defined as the ensemble of such points. The n variables are called Co-ordinates of the manifold.

With each point $P(u^1, u^2, \ldots, u^n)$ of the manifold we suppose associated a fundamental differential quadratic form:

$$(1) \qquad \overline{ds}^2 = g_{ij}du^i du^j;$$

ds is called a Line Element and represents the Distance between the point P and the infinitely near point $P'(u^1 + du^1, u^2 + du^2, \ldots, u^n + du^n)$, and is supposed invariant in any change of co-ordinates; the g-coefficients are given functions of the co-ordinates, with $g_{ji} = g_{ij}$, these functions, together with their first and second derivatives, being supposed finite and continuous.

The differential quadratic form will be assumed, unless otherwise stated, to be positive definite (i.e., positive for all values of the du's unless these are all zero); the distance between any two real points of the manifold will then be real, and the determinant of the g-coefficients of the form will be positive.

It will appear later that all the metrical properties of an n-dimensional manifold are determined through its associated differential quadratic form. It is, therefore, called a Metrical n-dimensional Manifold. It is also commonly called a Riemannian n-Space, after Riemann. In what follows we shall usually refer to an n-dimensional metrical manifold as a V_n.

The Direction of the point P' from the point P is determined by the differences of corresponding co-ordinates of the two points, that is, by the n du's, as in the case of a Euclidean V_3-space. The n quantities,

$$(2) \qquad \frac{du^1}{ds}, \frac{du^2}{ds}, \ldots, \frac{du^n}{ds},$$

are called the Direction Parameters of this direction. In virtue of equation (1) these parameters are subject to a relation analogous to that for the directions cosines of a direction in a Euclidean V_3. If the position of P' with respect to P is given, the direction parameters of the direction from P to P' are uniquely determined, and vice versa.

A continuous succession of line elements such as ds, constitutes what is called a Curve in the manifold. Such a curve can be specified by n parametric equations:

$$u^i = u^i(x), \qquad (i = 1, 2, \ldots, n),$$

where x is a continuous variable parameter; for a given value of x these equations determine a point on the curve. If, for a particular curve, all the u-functions are constant except u^1, then the curve is such that along it only the co-ordinate u^1 varies, and the curve is called a u^1-Curve. There will be a u-curve for each of the n co-ordinates.

If ds be an element of the u^1-curve, for example, at the point P, then the direction parameters of ds will each have the value zero except du^1/ds.

The set of direction parameters (2) for ds together with the magnitude of ds, determined by the differential quadratic form (1), is said to define an infinitesimal vector $d\mathbf{s}$ of magnitude ds, associated with the point P. Evidently, $d\mathbf{s}$ is an invariant for all co-ordinate systems.

More generally, a set of n direction parameters at a point P of a V_n, together with any positive scalar quantity A, determines a quantity $\mathbf{A}$, which is called a Vector at P of magnitude A[1].

§91

The Angular Metric and the Scalar Product of Two Vectors in a V_n

As we have seen, the line element ds given by the fundamental differential quadratic form (1), Art. 90, furnishes the basis of length measurement in a V_n, but as yet we have found no basis of angular measure. We have then to consider now the matter of the definition of the angle between two directions at a point P of a V_n.

For the special case of a V_2 we already have in formula (10), Art. 31, viz:

$$\cos(d\mathbf{s}, \delta\mathbf{s}) = g_{ij}\frac{du^i}{ds}\frac{\delta u^i}{\delta s}, \qquad (i, j = 1, 2),$$

an expression for the cosine of the angle between the two infinitesimal surface vectors $d\mathbf{s}$ and $\delta\mathbf{s}$. We now generalize this formula so as to apply to a V_n, by writing:[2]

$$\cos(d\mathbf{s}, \delta\mathbf{s}) = g_{ij}\frac{du^i}{ds}\frac{\delta u^j}{\delta s}, \qquad (i, j = 1, 2, \ldots, n), \tag{1}$$

[1] The vector here defined constitutes an n-fold hypernumber in the sense of Grassmann, mentioned in the historical introduction to the present book.

[2] It can be shown that the values which cos $(d\mathbf{s}, \delta\mathbf{s})$ can assume by this equation are all real, and lie within the limits $+1$ and -1 inclusive. See The Absolute Differential Calculus, by Levi Civita, Eng. tr., p. 123.

where $d\mathbf{s}$ and $\delta\mathbf{s}$ now denote infinitesimal vectors associated with a point P of the V_n, each of which determines a definite direction at the point P. Formula (1), with the convention that the cosines of the angles between like directions and between opposite directions are ± 1, furnishes a basis for the angular metric of the V_n.

The scalar product of the two infinitesimal vectors $d\mathbf{s}$ and $\delta\mathbf{s}$ is defined by the equation:

$$(2) \qquad d\mathbf{s} \cdot \delta\mathbf{s} = ds\delta s \cos (d\mathbf{s}, \delta\mathbf{s}).$$

From equations (1) and (2) we obtain the following bilinear differential form:

$$(3) \qquad d\mathbf{s} \cdot \delta\mathbf{s} = g_{ij}du^i\delta u^j.$$

Suppose now that $\mathbf{A}$ and $\mathbf{C}$, respectively, denote two vectors associated with the point P which have the arbitrary directions determined by $d\mathbf{s}$ and $\delta\mathbf{s}$, respectively, then the scalar product of these vectors is *defined* by the equation:

$$(4) \qquad \mathbf{A} \cdot \mathbf{C} = AC \cos (\mathbf{A}, \mathbf{C});$$

and du^i/ds and $\delta u^j/\delta s$ will be, respectively, typical direction parameters of $\mathbf{A}$ and $\mathbf{C}$, and we shall have by equation (1):

$$(5) \qquad \mathbf{A} \cdot \mathbf{C} = ACg_{ij}\frac{du^i}{ds}\frac{\delta u^j}{\delta s}.$$

Evidently $\mathbf{C} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{C}$, and scalar multiplication of two V_n vectors is therefore commutative.

If one of the vectors, say $\mathbf{C}$, is a unit vector, then the product $\mathbf{A} \cdot \mathbf{C}$ is called the Projection of $\mathbf{A}$ upon the direction determined by $\mathbf{C}$.

If neither of the vectors vanishes, and if $\mathbf{A} \cdot \mathbf{C} = 0$, then the vectors $\mathbf{A}$ and $\mathbf{C}$ are said to be mutually perpendicular.

§92

Unitary Vectors in a V_n

We now introduce a designation for the infinitesimal vector $d\mathbf{s}$ of Art. 90, which is suggested by the designations already used for an infinitesimal vector associated with a point in a Euclidean V_3, or with a point in a V_2, viz:

$$(1) \qquad d\mathbf{s} = \mathbf{a}_1du^1 + \mathbf{a}_2du^2 + \cdots + \mathbf{a}_ndu^n,$$

where the $\mathbf{a}$'s represent Unitary Vectors associated with the point $P(u^1, u^2, \ldots u^n)$ of a V_n, each of which is directed along the corresponding u-curve in the sense of u increasing.

From equation (1), Art 90, (3), Art. 91, and the last equation:

$$\overline{ds}^2 = \mathbf{ds} \cdot \mathbf{ds} = (\mathfrak{a}_1 du^1 + \mathfrak{a}_2 du^2 + \cdots + \mathfrak{a}_n du^n) \cdot (\mathfrak{a}_1 du^1 + \mathfrak{a}_2 du^2 + \cdots + \mathfrak{a}_n du^n) = (\mathfrak{a}_i du^i) \cdot (\mathfrak{a}_j du^j).$$

We now assume the distributive law to be valid for the indicated product on the right, and then find:

(2) $$\overline{ds}^2 = \mathfrak{a}_i \cdot \mathfrak{a}_j du^i du^j, \qquad \text{with } \mathfrak{a}_j \cdot \mathfrak{a}_i = \mathfrak{a}_i \cdot \mathfrak{a}_j.$$

Comparing this equation with equation (1), Art. 90, and noting that the du's are arbitrary, it appears that:

(3) $$\mathfrak{a}_i \cdot \mathfrak{a}_j = g_{ij},$$

and hence that:

(4) $$a_i a_j \cos\theta = g_{ij}, \qquad \text{with } \theta = (\mathfrak{a}_i, \mathfrak{a}_j),$$

where the a's represent magnitudes of the unitary vectors. By taking in this equation $j = i = 1, 2, \ldots n$, we then find for the magnitudes of the unitary vectors:

(5) $$a_1 = \sqrt{g_{11}}, \qquad a_2 = \sqrt{g_{22}}, \ldots\ldots, a_n = \sqrt{g_{nn}}.$$

The system of n unitary vectors associated with a point P of a V_n can, as exemplified by equation (1), be used as a base-system of reference in a manner quite analogous to that with which we are familiar in the case of a Euclidean V_3. Such a system will be called the U-system.

A vector **A** associated with the point P is designated on the U-system at P by writing[1]:

(6) $$\mathbf{A} = A^1\mathfrak{a}_1 + A^2\mathfrak{a}_2 + \cdots + A^n\mathfrak{a}_n.$$

The sum on the right is supposed subject to the same operational laws as a corresponding sum in a Euclidean V_3; the individual terms of the sum are called the Components of **A**, and the A-coefficients are called the Measure-Numbers of these components.

It should be noticed that all the formulas of the present article reduce, upon taking n equal to 3 or 2, to already familiar formulas for a Euclidean V_3 or for a V_2.

§93

Reciprocal Unitary Vectors in a V_n

Consider a system of n vectors, $\mathfrak{a}^1, \mathfrak{a}^2, \ldots, \mathfrak{a}^n$, associated with a point $P(u^1, u^2, \ldots, u^n)$ of a V_n. Suppose these vectors to be related to the unitary vectors associated with the same point as follows:

(1) $$\mathfrak{a}_i \cdot \mathfrak{a}^\lambda = g_i^\lambda,$$

[1] In accordance with the procedure of Grassmann.

where $g_i^\lambda = 1$, if $i = \lambda$, and $g_i^\lambda = 0$, if $i \neq \lambda$. The n $\mathfrak{a}^\lambda$-vectors which satisfy these relations will be called Reciprocal Unitary vectors. They also can be used as a base-system of reference for vectors associated with the point P, and this system will be called the R-system.

The infinitesimal position-vector $d\mathbf{s}$ of a point $P'(u^1 + du^1, u^2 + du^2, \ldots u^n + du^n)$ with respect to P is designated on the R-system, in analogy with the designation on the U-system expressed by equation (1), Art. 92, as follows:

(2) $$d\,\mathbf{s} = \mathfrak{a}^\lambda du_\lambda = \mathfrak{a}^\mu du_\mu,$$

where the du's are differentials which in general are non-integrable. From equations (2), by forming scalar products, we find:

(3) $$\overline{ds}^2 = \mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu du_\lambda du_\mu,$$

We now let:

(4) $$\mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu = g^{\lambda\mu}.$$

The preceding equation can then be written:

(5) $$\overline{\overline{ds}}^2 = g^{\lambda\mu} du_\lambda du_\mu, \qquad \text{with } g^{\lambda\mu} = g^{\mu\lambda}.$$

This expression for $\overline{\overline{ds}}^2$ will be called the Reciprocal Differential Quadratic Form for a V_n.

A vector $\mathbf{A}$ associated with the point P is designated on the R-system at P by writing:

(6) $$\mathbf{A} = A_1\mathfrak{a}^1 + A_2\mathfrak{a}^2 + \cdots + A_n\mathfrak{a}^n.$$

The sum on the right is supposed subject to the same operational laws as a corresponding sum in a Euclidean V_3; the individual terms are called the Components of $\mathbf{A}$ on the R-system, and the A-coefficients are called the Measure-Numbers of these components.

§94

Relationship of Unitary and Reciprocal Differentials and of Unitary and Reciprocal Unitary Vectors

From equations (1), Art. 92, and (2), Art. 93, we have:

$$\mathfrak{a}^\lambda du_\lambda = \mathfrak{a}_i du^i,$$

each member of this equation representing the fixed infinitesimal vector $d\mathbf{s}$. From this equation, by a procedure quite similar to

that used in deriving the analogous equations (3) and (4), Art. 74, for a Euclidean V_3, we find:

(1) $$du_\lambda = g_{\lambda i} du^i, \qquad du^i = g^{i\lambda} du_\lambda;$$

(2) $$\mathfrak{a}^\lambda = g^{\lambda i} \mathfrak{a}_i, \qquad \mathfrak{a}_i = g_{i\lambda} \mathfrak{a}^\lambda;$$

where:

(3) $$g_{\lambda i} = \mathfrak{a}_\lambda \cdot \mathfrak{a}_i = \mathfrak{a}_i \cdot \mathfrak{a}_\lambda = g_{i\lambda};$$

(4) $$g^{i\lambda} = \mathfrak{a}^i \cdot \mathfrak{a}^\lambda = \mathfrak{a}^\lambda \cdot \mathfrak{a}^i = g^{\lambda i}.$$

Equations (1) express the relationship of the unitary and reciprocal differentials, and equations (3) those of the unitary and reciprocal unitary vectors.

It appears from equations (2) that $g^{\lambda i}$ may be regarded as an operator which acting upon the unitary vector $\mathfrak{a}_i$ changes it into the reciprocal unitary vector $\mathfrak{a}^\lambda$, and that $g_{i\lambda}$ may be considered as an operator which acting upon the reciprocal unitary vector $\mathfrak{a}^\lambda$ changes it into the unitary vector $\mathfrak{a}_i$. These operations are examples, respectively, of the processes which were designated in Art. 74 as raising and lowering of indices.

By scalar multiplication of the second of equations (2) by $\mathfrak{a}^k$, taking account of equations (4), we find:

(5) $$g_{i\lambda} g^{k\lambda} = g_i^k,$$

where $g_i^k = 1$, if $k = i$, and $g_i^k = 0$, if $k \neq i$.

The determinants (g, g') of the coefficients of the two sets of equations for which the second and first of the linear equations (2) are typical can be expressed as follows:

$$|g_{11} \; . \; . \; . \; g_{1n}$$

(6) (7)

Solving, by the method of determinants, these two sets of equations, we find:

(8) $$\mathfrak{a}^\lambda = G^{i\lambda} \mathfrak{a}_i, \qquad \mathfrak{a}_i = G_{\lambda i} \mathfrak{a}^\lambda,$$

where $G^{i\lambda}$ is the cofactor of the element $g_{i\lambda}$ of the determinant g divided by g, and $G_{\lambda i}$ is the cofactor of the element $g^{\lambda i}$ of the determinant g' divided by g'. Upon comparison of equations (2) and (6), we see that:

(9) $$g^{\lambda i} = G^{i\lambda}, \qquad g_{i\lambda} = G_{\lambda i}.$$

From equations (1), which express the relationship of the measure-numbers of the components on the unitary and the reciprocal unitary base-systems of the infinitesimal vector $d\mathbf{s}$, it can be inferred that the relationship of the measure-numbers of the components of any vector $\mathbf{A}$ on these systems will be as follows:

$$(10) \qquad A_\lambda = g_{\lambda i} A^i, \qquad A^i = g^{i\lambda} A_\lambda,$$

where A^i and A_λ represent typical measure-numbers of the components of $\mathbf{A}$ on the unitary and reciprocal unitary base-systems respectively.

If in all the formulas of the present article we take $n = 3$, they reduce to corresponding formulas previously found for the special case of a Euclidean V_3.

§95

Transformation Equations

Let $v^1, v^2, \ldots v^n$ denote the co-ordinates on a second system of co-ordinates, which will be called the V-System, and suppose the u and v-co-ordinates for a generic point P of the V_n to be related in some way which is symbolically expressed by the following equations:

$$(1) \qquad v^l = v^l(u^1, u^2, \ldots u^n), \qquad u^i = u^i(v^1, v^2, \ldots v^n),$$

where the u-functions, and likewise the v-functions, are assumed mutually independent and differentiable.

Upon differentiation of these equations we obtain directly the equations of transformation for the differentials of the co-ordinates in passing from the U to the V-system, and vice versa:

$$(\mathrm{I}) \qquad dv^l = \frac{\partial v^l}{\partial u^i} du^i, \qquad du^i = \frac{\partial u^i}{\partial v^l} dv^l,$$

where the du's and the dv's may be regarded as the relative co-ordinates on the U and V-systems respectively of a point P' infinitely near to P.

The Jacobeans (J, K) of the two sets of equations for which equations (I) are typical can be expressed as follows:

$$(2) \qquad \begin{vmatrix} \dfrac{\partial v^1}{\partial u^1} & \dfrac{\partial v^1}{\partial u^n} \\ & \\ \dfrac{\partial v^n}{\partial u^1} & \dfrac{\partial v^n}{\partial u^n} \end{vmatrix} \qquad (3) \qquad K = \begin{vmatrix} \dfrac{\partial u^1}{\partial v^1} & \dfrac{\partial u^1}{\partial v^n} \\ & \\ \dfrac{\partial u^n}{\partial v^1} \cdots & \dfrac{\partial u^n}{\partial v^n} \end{vmatrix}$$

These determinants cannot vanish identically, since it is assumed that the u-co-ordinates are independent of one another, and likewise the v-co-ordinates. Furthermore, it can be shown, as in Art. 81 for a Euclidean V_3, that:

(4) $$\frac{\partial u^i}{\partial v^l}\frac{\partial v^l}{\partial u^k} = \delta_k,$$

where $\delta^i_k = 1$, *if* $k = i$; and $\delta^i_k = 0$, *if* $k \neq i$; and also, that:

(5) $$JK = 1.$$

Let $\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_n$ denote the unitary vectors on the V-system. If $d\mathbf{s}$ denote the position-vector of P' with respect to P, then $d\mathbf{s}$ can be designated on the U and V-systems respectively by writing:

(6) $$d\mathbf{s} = \mathfrak{a}_i du^i;$$

(7) $$d\mathbf{s} = \mathbf{b}_l dv^l.$$

The infinitesimal vector $d\mathbf{s}$ is a fixed vector and therefore an invariant. Hence:

$$\mathbf{b}_l dv^l = \mathfrak{a}_i du^i.$$

From this relation the transformation equations for the unitary vectors can easily be found with the aid of equations (I) above. We obtain directly:

$$\mathbf{b}_l dv^l = \mathfrak{a}_i \frac{\partial u^i}{\partial v^l}\, dv^l, \qquad \mathfrak{a}_i du^i = \mathbf{b}_l \frac{\partial v^l}{\partial u^i}\, du^i.$$

Since these equations must be valid for all possible values of the dv's and the du's, it follows that:

(II) $$\mathbf{b}_l = \frac{\partial u^i}{\partial v^l}\,\mathfrak{a}_i, \qquad \mathfrak{a}_i = \frac{\partial v^l}{\partial u^i}\,\mathbf{b}_l.$$

These are the transformation equations for the unitary vectors.

From equations (II) we find directly:

(9) $$\frac{\partial u^i}{\partial v^l} = \mathbf{b}_l \cdot \mathfrak{a}^i;$$

(10) $$\frac{\partial v^l}{\partial u^i} = \mathfrak{a}_i \cdot \mathbf{b}^l.$$

Upon comparison of equations (I) and (II) above with the corresponding equations (I) and (II), Art. 81, for the transformation of the differentials of general co-ordinates and for the unitary vectors in a Euclidean V_3, it is seen that the corresponding equations are the same, except that the indices in the transformations for a

general V_n cover the range 1, 2, . . . n, while those for a Euclidean V_3 cover the range 1, 2, 3.

It is hardly necessary to prove that all the other transformation equations which were obtained in Chapter VIII for a Euclidean V_3 can be generalized so as to apply to a V_n by simply extending the range to be covered by the identifying indices from 1, 2, 3 to 1, 2, . . . n.

We thus obtain the following transformation equations for various quantities associated with a generic point in a V_n:

(I) $$dv^l = \frac{\partial v^l}{\partial u^i} du^i, \qquad du^i = \frac{\partial u^i}{\partial v^l} dv^l;$$

(II) $$\mathfrak{b}_l = \frac{\partial u^i}{\partial v^l} \mathfrak{a}_i, \qquad \mathfrak{a}_i = \frac{\partial v^l}{\partial u^i} \mathfrak{b}_l;$$

(III) $$dv_\rho = \frac{\partial u^\lambda}{\partial v^\rho} du_\lambda, \qquad du_\lambda = \frac{\partial v^\rho}{\partial u^\lambda} dv_\rho;$$

(IV) $$\mathfrak{b}^\rho = \frac{\partial v^\rho}{\partial u^\lambda} \mathfrak{a}^\lambda, \qquad \mathfrak{a}^\lambda = \frac{\partial u^\lambda}{\partial v^\rho} \mathfrak{b}^\rho;$$

(V) $$h_{lm} = \frac{\partial u^i}{\partial v^l} \frac{\partial u^j}{\partial v^m} g_{ij}, \qquad g_{ij} = \frac{\partial v^l}{\partial u^i} \frac{\partial v^m}{\partial u^j} h_{lm};$$

(VI) $$h^{\rho\sigma} = \frac{\partial v^\rho}{\partial u^\lambda} \frac{\partial v^\sigma}{\partial u^\mu} g^{\lambda\mu} \qquad g = \frac{\partial u^\lambda}{\partial v^\rho} \frac{\partial u^\mu}{\partial v^\sigma} h^{\rho\sigma};$$

(VII) $$h = K^2 g, \qquad g = J^2 h;$$

(VIII) $$h' = J^2 g', \qquad g' = K^2 h';$$

(IX) $$B^l = \frac{\partial v^l}{du^i} A^i, \qquad A^i = \frac{\partial u^i}{\partial v^l} B^l;$$

(X) $$B_\rho = \frac{\partial u^\lambda}{\partial v^\rho} A_\lambda, \qquad A_\lambda = \frac{\partial v^\rho}{\partial u^\lambda} B_\rho.$$

These are the transformation equations for the following quantities associated with a generic point of a V_n: I—the differentials of the co-ordinates; II—the unitary vectors; III—the reciprocal differentials; IV—the reciprocal unitary vectors; V—the coefficients of the fundamental differential quadratic form; VI—the coefficients of the reciprocal differential quadratic form; VII—the determinant of the coefficients of the fundamental differential quadratic form; VIII—the determinant of the reciprocal differential quadratic form; IX—the contravariant measure-numbers of a vector; X—the covariant measure-numbers of a vector.

The quantities in the above equations which transform by the same law (I) as the differentials of the co-ordinates are said to be singly contravariant, and those which transform by the same law

(II) as the unitary vectors are said to be singly covariant. Thus, the differentials of the co-ordinates, the reciprocal unitary vectors, and the measure-numbers of the components of a vector on the unitary base-system are singly contravariant quantities; and the unitary vectors, the reciprocal differentials, and the measure-numbers of a vector on the reciprocal unitary base-system are singly covariant quantities.

For obvious reasons the coefficients of the fundamental differential quadratic form are said to be doubly covariant, and the coefficients of the reciprocal differential quadratic form are said to be doubly contravariant.

The determinants of the g-coefficients, which transform by equations (VII) and (VIII), were it not for the factor J^2 or K^2, would be invariants.

As regards notation, it will be noticed that, as in Chapters VII and VIII, covariant quantities are designated by subscripts and contravariant quantities by superscripts.

Finally, in connection with the above transformation equations, the explanatory remarks on covariant and contravariant quantities made in Art. 82 are all applicable to the covariant and contravariant quantities which appear in these equations.

§96

Several Important V_n-Invariants

The scalar product of two vectors associated with a generic point P of a V_n is a scalar invariant, as will be evident upon inspection of equations 1 below.

Let **A** and **C** denote the two vectors. Each of them can be referred to the unitary and the reciprocal unitary base-systems associated with the point P by writing:

$$\mathbf{A} = A^i\mathbf{a}_i = A_\lambda\mathbf{a}^\lambda;$$
$$\mathbf{C} = C^j\mathbf{a}_j = C_\mu\mathbf{a}^\mu.$$

Making use of the relation (1), Art. 93, we then find:

$$(1) \qquad \begin{aligned} \mathbf{A}\cdot\mathbf{C} &= g_{ij}A^iC^j \\ &= A^iC_i \\ &= A_\lambda C^\lambda \\ &= g^{\lambda\mu}A_\lambda C_\mu. \end{aligned}$$

Another important scalar invariant associated with the point P

is the fundamental differential quadratic form itself. For this we have, with the aid of equations (1), Art. 94:

$$\begin{aligned} \overline{ds}^2 &= \mathbf{ds} \cdot \mathbf{ds} = g_{ij} du^i du^j \\ &= du^i du_i \\ &= g^{\lambda\mu} du_\lambda du_\mu. \end{aligned} \tag{2}$$

A fixed vector associated with the point P affords an example of a non-scalar invariant. Let the vector be denoted by $\mathbf{A}$. Then, with the aid of equations (10) and (2), Art. 94, we see that:

$$\begin{aligned} \mathbf{A} &= A^i \mathbf{a}_i \\ &= g^{i\lambda} A_\lambda \mathbf{a}_i \\ &= g_{i\lambda} A^i \mathbf{a}^\lambda \\ &= A_\lambda \mathbf{a}^\lambda. \end{aligned} \tag{3}$$

§97

Geodetic Lines in a V_n

Let P_1 and P_2 be any two points in a V_n, and let l denote the length of a curve connecting the two points, and ds an element of this curve. Then:

$$l = \int_{P_1}^{P_2} ds.$$

If $l + \delta l$ be the length of a second curve connecting the points P_1 and P_2 which lies infinitely near to the first, and if

$$\delta l \quad \delta \int_{P_1}^{P_2} ds = 0,$$

the original curve is called a Geodetic Line. By this definition, the length of a geodetic line must be an extremum as regards infinitely near curves connecting the same two points.

The problem of finding geodetic lines is analogous to that of finding maximum and minimum values of algebraic functions, but is more complicated owing to the fact that we have to seek for extremum lines instead of extremum points.

The points which lie on a geodetic line must be such that their co-ordinates satisfy certain differential equations which will be derived presently. The process of derivation of these equations is quite the same whatever the value of n; geometrical visualization of the process is, however, only possible (except in the special case of a Euclidean V_3) in the case for which $n = 2$, that is, the case of an ordinary surface. We shall therefore restrict ourselves to this case, arriving finally at two differential equations for a geodetic

line, from which the corresponding equations for a geodetic line in a general V_n can be directly inferred.

Accordingly, we consider a curve drawn on any surface S connecting two points P_1 and P_2, of length

$$l \quad \int_{P_1}^{P_2} d$$

and a second curve, drawn on the surface between the same two points P_1 and P_2, whose length differs from l by the amount δl. The two curves in question are represented in Fig. 40. The distance measured along the first curve (g) from the initial point P_1 to any other point $P(u^1, u^2)$ is denoted by s, and $s + ds$ denotes the corresponding distance for the neighboring point $Q(u^1 + du^1, u^2 + du^2)$, u^1 and u^2 denoting parametric co-ordinates of the surface. The corresponding primed quantities s' and $s' + ds'$ have a similar significance as regards the second line. The point P' is, by supposition, infinitely close to P, and it is further supposed that the first line is subdivided into infinitesimal elements of which ds is typical, and the second into corresponding elements of which ds' is typical, the correspondence being one to one.

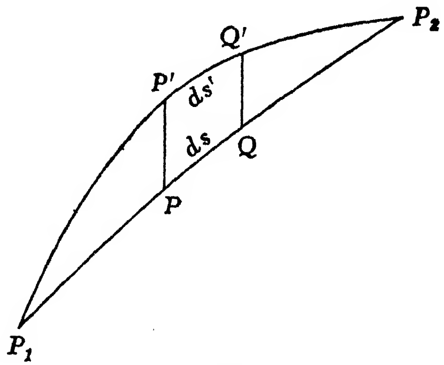


Fig. 40.

The co-ordinates u^1, u^2 of P are, of course, functions of the parameter s; and the line g is defined by these functions which, as we shall see, must satisfy certain ordinary characteristic differential equations of the second order.

The symbol δ will be used to denote differences in corresponding quantities relating to the two lines. Two relations involving the symbol δ are easily derived as follows:

$$\delta ds = ds' - ds = d(s' - s) = d\delta s,$$

showing that the differential operators d and δ are commutative;

$$\delta(ds)^2 = (ds')^2 - (ds)^2 = (ds' + ds)\,(ds' - ds) = 2ds\delta s,$$

neglecting a small quantity of higher order. We shall have use presently for both of these relations.

The differential quadratic form for the surface can be expressed as follows:[1)]

$$\overline{ds}^2 = g_{ij}du^i du^j,$$

where, of course, the g-coefficients are in general functions of the co-ordinates u^1 and u^2. Operating with δ upon both sides of this equation, we get:

$$2ds\delta ds = \delta g_{ij}du^i du^j + g_{ij}\delta du^i du^j + g_{ij}du^i \delta du^j.$$

But:

$$\delta g_{ij} = \frac{\partial g_{ij}}{\partial u^k}\delta u^k,$$

and hence:

$$\delta g_{ij}du^i du^j = \frac{\partial g_{ij}}{\partial u^k}du^i du^j \delta u^k = \frac{\partial g_{ik}}{\partial u^j}du^i du^k \delta u^j,$$

since j and k are dummy indices and can, therefore, be interchanged. Hence, after division by $2ds$, the preceding equation for $2ds\delta ds$ can be written:

$$\delta ds = \frac{1}{2}\frac{\partial g_{ik}}{\partial u^j}\frac{du^i}{ds}\frac{du^k}{ds}\delta u^j ds + \frac{1}{2}g_{ij}\frac{du^j}{ds}d\delta u^i + \frac{1}{2}g_{ij}\frac{du^i}{ds}d\delta u^j.$$

Now:

$$\delta l = \delta\int_{P_1}^{P_2} ds = \int_{P_1}^{P_2}\delta ds.$$

Making use of the preceding equation, we find:

$$\delta l = \int_{P_1}^{P_2}\frac{1}{2}\frac{\partial g_{ik}}{\partial u^j}\frac{du^i}{ds}\frac{du^k}{ds}\delta u^j ds + J,$$

where:

$$J = \int_{P_1}^{P_2}\frac{1}{2}\left[g_{ij}\frac{du^j}{ds}d\delta u^i + g_{ij}\frac{du^i}{ds}d\delta u^j\right]$$

$$= \int_{P_1}^{P_2} g_{ij}\frac{du^i}{ds}d\delta u^j,$$

noting that, by the dummy index rule and the relation $g_{ij} = g_{ji}$, the two terms in brackets are equal. After integration by parts we have:

$$\left[g_{ij}\frac{du^i}{ds}\delta u^j\right]_{P_1}^{P_2} - \int_{P_1}^{P_2} d\left(g_{ij}\frac{du^i}{ds}\right)\delta u^j.$$

[1)] In this discussion it is to be understood that all identifying indices cover the range 1, 2.

The integrated part on the right must vanish, since at P_1 and P_2 the δu's must all vanish; and the integral can be split up into two, so that:

$$J = -\int_{P_1}^{P_2} g_{ij}\frac{d^2u^i}{ds^2}\,\delta u^j ds - \int_{P_1}^{P_2} dg_{ij}\frac{du^i}{ds}\,\delta u^j.$$

For the integrand of the second integral we can write:

$$\frac{1}{2}\left[\frac{\partial g_{ij}}{\partial u^k}\,du^k\frac{du^i}{ds}\,\delta u^j + \frac{\partial g_{kj}}{\partial u^i}\,du^i\frac{du^k}{ds}\,\delta u^j\right],$$

noting that the second term enclosed by brackets is equal to the first, since i and k are dummy indices, which can be interchanged. It follows, then, that:

$$J = -\int_{P_1}^{P_2} g_{ij}\frac{d^2u^i}{ds^2}\,\delta u^j ds - \int_{P_1}^{P_2}\frac{1}{2}\left[\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{kj}}{\partial u^i}\right]\frac{du^i}{ds}\frac{du^k}{ds}\,\delta u^j ds.$$

Upon insertion of this expression for J in the last equation for δl we get:

$$\delta l = -\int_{P_1}^{P_2}\left[-\frac{1}{2}\frac{\partial g_{ik}}{\partial u^j}\frac{du^i}{ds}\frac{du^k}{ds} + g_{ij}\frac{d^2u^i}{ds^2} + \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{kj}}{\partial u^i}\right)\frac{du^i}{ds}\frac{du^k}{ds}\right]\delta u^j ds,$$

or, after some rearrangement of terms in the integrand of the integral on the right:

$$\delta l = -\int_{P_1}^{P_2}\left[\frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j}\right)\frac{du^i}{ds}\frac{du^k}{ds} + g_{ij}\frac{d^2u^i}{ds^2}\right]\delta u^j ds.$$

Now, let:

(1) $$[ik, j] = \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j}\right), \qquad (i, j, k = 1, 2);$$

the symbol on the left is known as a Christoffel 3-Index Symbol of the first kind. We then have:

$$\delta l = -\int_{P_1}^{P_2}\left\{[ik, j]\frac{du^i}{ds}\frac{du^k}{ds} + g_{ij}\frac{d^2u^i}{ds^2}\right\}\delta u^j ds.$$

If the curve g is to be a geodetic, δl must vanish, and, on account of the arbitrary nature of the δu's, it is therefore necessary that:

(2) $$g_{ij}\frac{d^2u^i}{ds^2} + [ik, j]\frac{du^i}{ds}\frac{du^k}{ds} = 0.$$

Since the index j can assume each of the values 1 and 2, we have here two ordinary differential equations of the second order. It will be noticed that summation with respect to the index i is demanded in the first term and with respect to both of the indices i and k in the second term.

Multiplying equation (2) by g^{lj} and taking account of the relation (5), Art. 94, we find that the differential equations of a geodetic line in a V_2 can be written as follows:

$$(3) \qquad \frac{d^2u^l}{ds^2} + \{ik, l\} \frac{du^i}{ds} \frac{du^k}{ds} = 0, \qquad (i, j, k, l = 1, 2),$$

where:

$$(4) \qquad \{ik, l\} = g^{lj}[ik, j], \qquad (i, j, k, l = 1, 2).$$

The symbol on the left is called a Christoffel 3-Index Symbol of the second kind.

The process followed above in finding the differential equations of a geodetic line in a V_2 can be extended without adding anything new by way of principle so as to yield the differential equations of a geodetic line in a V_n. The result, as might be anticipated by analogy, is to give for the differential equations for a geodetic line in a V_n:

$$(5) \qquad \frac{d^2u^l}{ds^2} + \{ik, l\} \frac{du^i}{ds} \frac{du^k}{ds} = 0, \qquad (i, j, k, l = 1, 2 \ldots n),$$

where:

$$(6) \qquad \{ik, l\} = g^{lj}[ik, j], \qquad (i, j, k, l = 1, 2 \ldots n),$$

where:

$$(7) \qquad [ik, j] = \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j}\right), \qquad (i, j, k = 1, 2 \ldots n).$$

The n differential equations (5) of the second order are satisfied by the co-ordinates of any point on a geodetic line connecting any two points of a V_n. When integrated they give the parametric equations of the geodetic line. The $2n$ constants of integration can be determined when the co-ordinates of the two points are given, or when the co-ordinates of one of the points and the direction of the line at that point are specified. It should be noticed that these equations involve no quantities which are not intrinsic to the V_n.

Properties of Christoffel 3-Index Symbols

Equations (7) and (6) of the preceding article, viz.:

$$(a) \qquad [ik, j] = \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j}\right), \qquad (i, j, k = 1, 2, \ldots n),$$

$$(b) \qquad \{ik, l\} = g^{lj}[ik, j], \qquad (i, j, k, l = 1, 2, \ldots n),$$

can be considered as the defining equations for the Christoffel symbols of the first and second kinds, respectively, for a general V_n. There are evidently n symbols of each kind for each independent g_{ik} and, since the number of independent g_{ik}'s is not n^2 but $\frac{n(n+1)}{2}$, on account of the relations $g_{ik} = g_{ki}$, the number of independent Christoffel symbols of each kind is $\frac{n^2(n+1)}{2}$.

It is evident that the Christoffel symbols of both kinds are symmetric with respect to the indices i and k. Hence:

$$[ik, j] = [ki, j]; \tag{1}$$

$$\{ik, l\} = \{ki, l\}. \tag{2}$$

Upon multiplication of equation (b) by g_{lm} and taking into account the relation (5), Art. 94, we find:

$$[ik, j] = g_{lj}\{ik, l\}. \tag{3}$$

With the aid of equation (b) we see that:

$$\{ik, l\}[\alpha\beta, l] = g^{lj}[ik, j][\alpha\beta, l] = [ik, j]\{\alpha\beta, j\},$$

and, upon replacing the dummy index j by the dummy index l, we shall therefore have:

$$\{ik, l\}[\alpha\beta, l] = [ik, l]\{\alpha\beta, l\}. \tag{4}$$

From equation (a) the following equation can easily be derived:

$$\frac{\partial g_{ij}}{\partial u^m} = [im, j] + [jm, i], \qquad \text{identically}; \tag{5}$$

or, with the aid of equation (3) above:

$$\frac{\partial g_{ij}}{\partial u^m} = g_{lj}\{im, l\} + g_{li}\{jm, l\}, \qquad \text{identically}. \tag{6}$$

From the last equation, by making use of the relation (5), Art. 94, the following equation is easily found:

$$\frac{\partial g^{\lambda\mu}}{\partial u^m} = -g^{\rho\mu}\{\rho m, \lambda\} - g^{\rho\lambda}\{\rho m, \mu\}, \qquad \text{identically}. \tag{7}$$

The number of independent equations of the type (5) would be n^3 were it not for the $\frac{n(n-1)}{2}$ relations, $g_{ij} = g_{ji}$. Upon taking account of these relations, the number of independent equations is reduced to $\frac{n^2(n+1)}{2}$. This is also, as noted above, the number

of independent Christoffel symbols of either kind in a general V_n. Equations (5) are, therefore, just sufficient in number to determine these symbols in terms of the first derivatives with respect to the co-ordinates of the g-coefficients of the fundamental differential quadratic form upon which the metrical properties of a V_n depend. When, therefore, the metrical properties of the space are known, through specification of the g-coefficients of its fundamental differential quadratic form, the Christoffel symbols of both kinds will also be known.

An important formula expressing the derivative of the logarithm of the square root of the fundamental metrical determinant;

$$g =$$

$$|g_{n1}$$

can easily be derived by a method which will be exemplified for the special case of a V_2.

For this case we have:

$$g = \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix}$$

where the g's are functions of the co-ordinates u^1 and u^2. By the rules for the differentiation of a determinant:

$$\frac{\partial g}{\partial u^m} = \begin{vmatrix} \frac{\partial g_{11}}{\partial u^m} & \frac{\partial g_{12}}{\partial u^m} \\ g_{21} & g_{22} \end{vmatrix} + \begin{vmatrix} g_{11} & g_{12} \\ \frac{\partial g_{21}}{\partial u^m} & \frac{\partial g_{22}}{\partial u^m} \end{vmatrix}$$

$$\frac{\partial g_{11}}{\partial u^m} g_{22} - \frac{\partial g_{12}}{\partial u^m} g_{21} + \frac{\partial g_{22}}{\partial u^m} g_{11} - \frac{\partial g_{21}}{\partial u^m} g_{12}$$

$$= \frac{\partial g_{11}}{\partial u^m} G^{11} g + \frac{\partial g_{12}}{\partial u^m} G^{12} g + \frac{\partial g_{21}}{\partial u^m} G^{21} g + \frac{\partial g_{22}}{\partial u^m} G^{22} g,$$

where $G^{i\lambda}$ (with $i, \lambda = 1, 2$) represents the cofactor of the element $g_{i\lambda}$ in the determinant g, obtained by deleting the i row and the λ column, divided by g. Now, by the first of equations (9), Art. 94:

$$G^{i\lambda} = g^{\lambda i} = g^{i\lambda}, \qquad (i, \lambda = 1, 2).$$

Hence, making use of the summation convention, we can write:

$$\frac{\partial g}{\partial u^m} = \frac{\partial g_{i\lambda}}{\partial u^m} g^{i\lambda} g,$$

or, after division by g:

$$\frac{\partial}{\partial u^m} \log_e g = \frac{\partial g_{i\lambda}}{\partial u^m} g^{i\lambda}.$$

or, taking account of equation (6) above:

$$\frac{\partial}{\partial u^m} \log_e g = (g_{ki}\{\lambda m, k\} + g_{k\lambda}\{im, k\})\, g^{i\lambda}$$
$$= \{km, k\} + \{km, k\} = 2\{km, k\}, \qquad (i, k, \lambda, m = 1, 2).$$

Hence, after division by 2, we obtain finally:

$$\frac{\partial}{\partial u^m} \log_e \sqrt{g} = \{km, k\}, \qquad (k, m = 1, 2).$$

The process followed in arriving at this formula for a V_2 can be extended without difficulty to allow of the derivation of the corresponding formula for a V_n, viz:

(8) $$\frac{\partial}{\partial u^m} \log_e \sqrt{g} \quad \{km, k\}, \qquad (k, m = 1, 2, \ldots n).$$

§99

Parallel and Equipollent Vectors in a Surface

Two vectors, supposed localized, respectively, at two points P and P' of a Euclidean plane are parallel in the Euclidean sense, provided they make the same angle with the straight line connecting P and P'. This line is, of course, a geodetic line connecting the two points.

These facts suggest a method of procedure whereby the parallelism of two surface vectors localized, respectively, at two points of a developable surface can be defined. If **A** and **C** denote two surface vectors, localized respectively at the points P and P' of a developable surface S, we shall say that they are parallel, provided they make equal angles with the geodetic line on the surface connecting the points P and P'.

Since S is supposed a developable surface, such as that of a cylinder or of a cone, it can be rolled out on a plane without stretching, tearing, or overlapping, and in doing this lengths of lines and angles between lines will not be altered. In particular, the geodetic line connecting the points P and P' will, after development, be a straight line, and the two surface vectors **A** and **C** will, if originally parallel, be parallel in the Euclidean sense.

In the case of a non-developable surface S, such as that of a sphere, it is a well-known geometrical fact that a developable surface S' is imaginable which is tangent to the non-developable surface along any curve (T) connecting two points P and P'. If, then, **A** and **C** are two surface vectors, common to S and S', at the

points P and P', respectively, which make the same angle with the geodetic line (g') in S' connecting the points P and P', then these surface vectors will be parallel surface vectors for S', in accordance with the definition given above; furthermore, the vectors **A** and **C**, regarded as vectors in the surface S, are said (by Levi-Civita) to be parallel vectors *with respect to the curve* T.

Parallelism of vectors at two distant points on a non-developable surface S is thus only defined with respect to a surface curve T connecting the two points.

If the curve T is itself a geodetic line in S, it is evident at once that parallelism of two surface vectors is a property which depends only upon the intrinsic nature of the surface, and, therefore, upon the metrical coefficients of its fundamental differential quadratic form. It can be shown that this is also true when T is any surface curve.[1)]

Two surface vectors which are parallel with respect to a curve T and whose magnitudes are equal will be called Equipollent Surface Vectors with respect to the curve T.

§100

Absolute Differentiation

Consider a continuous single-valued surface-vector point function, and let its value at any point $P(u^1, u^2)$ of a surface S be denoted by **A**, and at the infinitely near point $P'(u^1 + du^1, u^2 + du^2)$ by **A**$'$. If $d\mathbf{A}$ be defined as an infinitesimal surface vector at P which is equal to the equipollent vector of **A**$'$ at P, with respect to the infinitesimal geodetic arc connecting P' and P, minus the vector **A**, then $d\mathbf{A}$ will be called the Absolute Differential of **A** in passing from P to P'.

The absolute differential of a scalar point function is assumed identical with its ordinary differential.

For the purposes which we now have in view it is necessary to define the significance of the absolute differential of a unitary vector at the point P in passing from P to P'.

In order to do this, we consider a unit surface-vector $d\mathbf{s}/ds$ which at all points along the geodetic line passing through the points P and P' is directed tangentially to this line. The absolute differential of this unit vector in passing from P to P' will be zero, its

[1)] For a proof of this statement the reader is referred to The Absolute Differential Calculus, by Levi-Civita, Eng. Tr., p. 106.

magnitude and direction remaining unchanged. The unit vector can be expressed in terms of its components on the unitary base-system at P by writing:

$$\frac{d\mathbf{s}}{ds} = \alpha_l \frac{du^l}{ds}, \qquad (l = 1, 2).$$

For its absolute differential we shall then have:

$$d\frac{d\mathbf{s}}{ds} \quad d\left(\alpha_l \frac{du^l}{ds}\right) = 0, \qquad (l = 1, 2),$$

where d is a symbol of absolute differentiation. We now assume that the absolute differentiation of the expression in parentheses is to follow the laws of ordinary differentiation. We can then write:

$$d\alpha_l \frac{du^l}{ds} + \alpha_l \frac{d^2u^l}{ds} \quad 0, \qquad (l = 1, 2).$$

The first factor of the first term on the left of this equation is called the absolute differential of the unitary vector α_l in passing from P to P'. It is the precise significance which such a differential shall have that we are now seeking to define.

We now assume that $d\alpha_i$ is a surface vector associated with the point P which, like any other vector associated with the same point, can be expressed in terms of its components on the unitary base-system at P. We can then write:

(1) $$d\alpha_i = \alpha^l \cdot d\alpha_i \alpha_l, \qquad (i, l = 1, 2).$$

Now, upon scalar multiplication of the equation preceding this by α^k/ds and using the dummy index rule, we find:

$$\frac{d^2u^l}{ds^2} + \frac{du^i}{ds}\alpha^l \cdot \frac{d\alpha_i}{ds} = 0, \qquad (i, l = 1, 2).$$

This equation must be compatible with equation (3), Art. 97, for a geodetic line, and will be under the assumption[1] which we now make, viz:

(2) $$\alpha^l \cdot d\alpha_i = \{ik, l\}\, du^k, \qquad (i, k, l = 1, 2),$$

where the bracketed expression represents a Christoffel 3-index symbol of the second kind. From equations (1) and (2) we now obtain:

(3) $$d\alpha_i = \alpha_l \{ik, l\}\, du^k, \qquad (i, l, k = 1, 2).$$

[1] This assumption is permissible if consistent with the identity equation (6), Art. 98. That this is the case is seen at once upon differentiation of $\alpha_i \cdot \alpha_j$ $(= g_{ij})$ with respect to u^k, making use of the relation $\alpha_i = g_{il}\alpha^l$, and introducing the assumption in question.

This equation gives the significance which is to be attached to the absolute differential $d\mathfrak{a}_i$ of the unitary vector $\mathfrak{a}_i$ in passing from the point $P(u^1, u^2)$ to the point $P'(u^1 + du^1, u^2 + du^2)$ of the surface.

It will now be shown that, corresponding to equation (3) for the absolute differential of a unitary vector, we shall have for the absolute differential of a reciprocal unitary vector at the point P of the surface the following equation:

$$(4) \qquad d\mathfrak{a}^\lambda = -\mathfrak{a}^\rho \{\rho k, \lambda\}\, du^k, \qquad (\lambda, k, \rho = 1, 2).$$

By equation (1), Art. 93, we have:

$$\mathfrak{a}_\rho \cdot \mathfrak{a}^\lambda = g^\lambda_\rho, \qquad (\lambda, \rho = 1, 2),$$

where $g^\lambda_\rho = 1$, if $\lambda = \rho$, and $g^\lambda_\rho = 0$, if $\lambda \neq \rho$. In either case, upon absolute differentiation of this equation, under the assumption that absolute differentiation of the scalar product $\mathfrak{a}_\rho \cdot \mathfrak{a}^\lambda$ is effected in the same manner as in ordinary differentiation, we find:

$$d(\mathfrak{a}_\rho \cdot \mathfrak{a}^\lambda) = \mathfrak{a}_\rho \cdot d\mathfrak{a}^\lambda + \mathfrak{a}^\lambda \cdot d\mathfrak{a}_\rho = 0, \qquad (\lambda, \rho = 1, 2).$$

Hence, with the aid of equation (2), we get:

$$(5) \qquad \mathfrak{a}_\rho \cdot d\mathfrak{a}^\lambda = -\{\rho k, \lambda\}\, du^k, \qquad (\lambda, k, \rho = 1, 2).$$

We now assume that $d\mathfrak{a}^\lambda$, like $d\mathfrak{a}_i$, is a vector associated with the point P; we can then write:

$$d\mathfrak{a}^\lambda = \mathfrak{a}_\rho \cdot d\mathfrak{a}^\lambda \mathfrak{a}^\rho, \qquad (\lambda, \rho = 1, 2).$$

From the last two equations the validity of equation (4) follows at once.

By analogy with equations (3) and (4) for a V_2 we now define the absolute differentials of unitary and reciprocal unitary vectors in a general V_n as follows:

$$(6) \qquad d\mathfrak{a}_i = \mathfrak{a}_l \{ik, l\}\, du^k, \qquad (i, k, l = 1, 2, \ldots n);$$

$$(7) \qquad d\mathfrak{a}^\lambda = -\mathfrak{a}^\rho \{\rho k, \lambda\}\, du^k, \quad (\lambda, k, \rho = 1, 2, \ldots n).$$

These equations differ from those for a V_2 only in that the indices have a range of integer values from 1 to n instead of from 1 to 2. The following expressions for the partial derivatives of unitary and reciprocal unitary vectors are derivable from these equations:

$$(8) \qquad \frac{\partial \mathfrak{a}_i}{\partial u^k} = \{ik, l\}\, \mathfrak{a}_l, \qquad (i, k, l = 1, 2, \ldots n),$$

$$(9) \qquad \frac{\partial \mathfrak{a}^\lambda}{\partial u^k} = -\{\rho k, \lambda\}\, \mathfrak{a}^\rho, \qquad (\lambda, k, \rho = 1, 2, \ldots n).$$

Next, we consider the case of any vector function **A** at a point P in a general V_n. Let it be expressed in the alternative forms:

(10) $\mathbf{A} = A^i \mathfrak{a}_i$; (11) $\mathbf{A} = A_\lambda \mathfrak{a}^\lambda$, $(i, \lambda = 1, 2, \ldots n)$;

where A^i, A_λ are, respectively, contravariant and covariant measure-numbers of the components of **A**. Upon absolute differentiation of these forms, making use of the dummy index rule, we get:

$$d\mathbf{A} = \mathfrak{a}_l dA^l + A^i d\mathfrak{a}_i, \qquad (i, l = 1, 2, \ldots n);$$
$$d\mathbf{A} = \mathfrak{a}^\rho dA_\rho + A_\lambda d\mathfrak{a}^\lambda, \qquad (\lambda, \rho = 1, 2, \ldots n);$$

or, with the aid of equations (6) and (7):

$$(12) \qquad d\mathbf{A} = [dA^l + A^i\{ik, l\}\, du^k]\, \mathfrak{a}_l, \qquad (i, k, l = 1, 2, \ldots n);$$
$$(13) \qquad d\mathbf{A} = [dA_\rho - A_\lambda\{\rho k, \lambda\}\, du^k]\, \mathfrak{a}^\rho, \qquad (\lambda, k, \rho = 1, 2, \ldots n).$$

These are alternative forms for the absolute differential of a vector in a general V_n; the cofactor of $\mathfrak{a}_l$ is a contravariant component and the cofactor of $\mathfrak{a}^\rho$ is a covariant component of $d\mathbf{A}$; but the individual terms in the components belong, in general, neither to a contravariant nor to a covariant system.

§101

Parallel Displacement of a Vector around a Spherical Polygon

By the term *parallel displacement of a vector in a surface S with respect to a path T* is to be understood a displacement whereby the point with which the vector is associated moves along T while the direction of the vector is continuously modified so that, if $\mathbf{A}_0$ denote its value at the initial point P_0 and $\mathbf{A}'$ its value at any other point P' of T, then $\mathbf{A}'$ and $\mathbf{A}_0$ will be equipollent[1] with respect to the path P_0P'.

The magnitude of a surface vector is, of course, unchanged by parallel displacement.

Furthermore, the angle between two surface vectors associated with the same point obviously remains unaltered by their parallel displacement with respect to the same path.

We shall now determine the change in a surface vector **A** associated with a spherical surface S produced by parallel displacement along the (geodetic) sides of a spherical triangle, or any n-sided

[1] For the definition of an equipollent vector, see the last paragraph of Art. 99.

spherical polygon, and we shall find that it bears a certain relation to the spherical excess[1)] of the triangle or polygon.

Since in the parallel displacement of a vector its magnitude remains unaltered, the change, if any, undergone by $\mathbf{A}$ in the supposed displacement must be a change of direction only. We shall see that there is such a change, and, in fact, that the angular change in direction of $\mathbf{A}$ in a parallel displacement around the triangle or polygon is measured by the corresponding spherical excess.

Referring to Fig. 41, ABC represents a triangle on the surface of a sphere, and $\underline{\mathfrak{a}}_1$ and $\underline{\mathfrak{a}}_2$ line-vectors representing the unitary vectors at the point A for which the Gaussian co-ordinates are u^1 and u^2, $\underline{\mathfrak{a}}_1$ and $\underline{\mathfrak{a}}_2$ being, of course, drawn tangentially to the u^1 and u^2-curves respectively in the directions of u^1 and u^2 increasing. The line-vector $\underline{\mathbf{A}}'$ represents the vector $\mathbf{A}'$ which is produced in a parallel displacement of the vector $\mathbf{A}$, once around the spherical triangle ABC in the direction of circulation which is definitionally taken as positive, viz: $A \rightarrow B \rightarrow C$ or, in other words, the direction in which $\mathfrak{a}_1$ must be rotated in order to *decrease* the angle between $\underline{\mathfrak{a}}_1$ and $\underline{\mathfrak{a}}_2$. The angles made by $\mathbf{A}$ and $\mathbf{A}'$ with the base-vector $\mathfrak{a}_1$ are denoted by α and α', respectively.

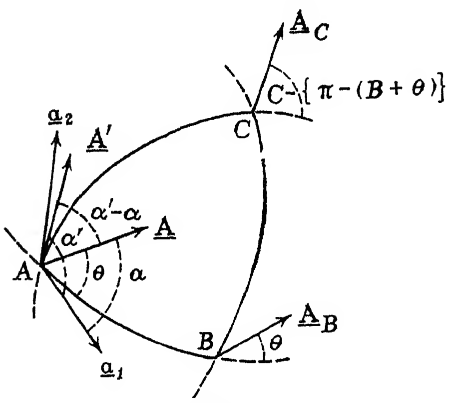


Fig. 41.

Now, when $\mathbf{A}$ undergoes a parallel displacement from A to B, since it continues to make the same angle, θ say, with the geodetic AB, the angle made at B by $\mathbf{A}$ with the geodetic BC will be $\pi - (B + \theta)$; in the further parallel displacement of $\mathbf{A}$ along the geodetic from B to C, the vector will continue to make the angle $\pi - (B + \theta)$ with this geodetic, and, therefore, the angle made by the vector $\mathbf{A}$ at C with the geodetic AC will be $C - \{\pi - (B + \theta)\}$ or $B + C + \theta - \pi$; in the final stage of the parallel displacement of the vector

[1)] The spherical excess of an n-sided spherical polygon is equal to the sum of its interior angles diminished by the angle $(n - 2)\pi$; in the special case of a spherical triangle, $n = 3$, of course.

A from C back to A along the geodetic AC, it will continue to make this angle with AC, and at A becomes the vector which we have designated by **A′**. From the diagram in Fig. 41, it is, then, at once evident that:

$$\alpha' - \alpha = A + B + C - \pi. \tag{1}$$

The expression on the right represents the spherical excess (ϵ) of the spherical triangle ABC, and, consequently:

$$\alpha' - \alpha = \epsilon. \tag{2}$$

The process used above in finding the change of direction produced in a vector in its parallel displacement around a spherical triangle can also be used to find the change of direction ($\alpha' - \alpha$) produced in its parallel displacement around any n-sided spherical polygon[1]; in this case the process yields the following result:

$$\alpha' - \alpha = \sum_n A_i - (n - 2)\,\pi, \tag{3}$$

where the first term on the right denotes the sum of the interior angles of the polygon. The entire right-hand member of this equation represents the spherical excess, ϵ say, of the polygon and, as in the case of the spherical triangle, which is, of course, a special case of the polygon, we have:

$$\alpha' - \alpha = \epsilon. \tag{4}$$

It should be noticed that the results here found are independent of the size of the polygon and, in virtue of symmetry, of its position on the sphere.

By an elementary theorem of solid geometry:

$$\frac{\sigma}{R^2}, \tag{5}$$

where σ represents the area of an n-sided polygon[1] upon a sphere of radius R. Furthermore, the Gaussian curvature K of a sphere of radius R by equation (30), Art. 31, is given by the equation:

$$K = \frac{1}{R^2}. \tag{6}$$

Accordingly, we can write:

$$K = \frac{\epsilon}{\sigma} \tag{7}$$

or, taking account of equation (4):

$$K = \tag{8}$$

[1] It is to be understood that the sides are parts of geodetic lines.

§102

Gaussian Curvature of a Surface—Riemannian Curvature of a V_n

By equation (7) of the preceding article the following rule is seen to be valid:

The Gaussian curvature of any spherical surface multiplied by the area of any n-sided polygon[1] supposed drawn on the surface is equal to the spherical excess of the polygon.

There is a corresponding rule which is valid in the case of a surface in general, and which can be stated as follows:

The *surface integral* of the Gaussian curvature over the area of any n-sided polygon[1] supposed drawn on the surface is equal to the angular excess of the sum of the interior angles of the polygon over the sum of the interior angles of on ordinary n-sided plane polygon.

This would appear to be the probable generalization of the corresponding rule stated above for spherical surfaces. The actual proof will not be given here; it follows directly from a celebrated theorem of the differential geometry of surfaces, known as the Integral Formula of Gauss-Bonnet.[2]

Corresponding to formula (8) of the preceding article for the case of a spherical surface, we can now write the following integral formula applicable to a surface in general:

$$\int K d\sigma = \alpha' - \alpha, \tag{1}$$

where $\alpha' - \alpha$ represents the angular change produced in a surface-vector by its parallel displacement once around an n-sided[1] polygon supposed drawn on the surface, the change being reckoned as positive when in the direction of displacement of the vector around the polygon.

By formula (1) we can now write for the Gaussian curvature at a generic point on a general surface:

$$K = \frac{D\alpha}{D\sigma}, \tag{2}$$

where $D\sigma$ represents the area of an infinitesimal element of the surface bounded by a contour line, consisting of elements of geodetic lines, which passes through the point in question, and $D\alpha$ represents

[1] It is to be understood that the sides are parts of geodetic lines.
[2] Cf. W. Blaschke, Differentialgeometrie, Art. 63.

the infinitesimal angular change in the parallel displacement of a surface-vector once around this contour.

By means of the definition of the Gaussian curvature K of a V_2 given above, the Riemannian curvature of a general V_n with respect to a section at any point P can be defined as follows:

Any two vectors **A**, **C** at a generic point P of a general V_n determine a section of the V_n with ∞^1 directions, to each of which there will correspond direction parameters, two in number, with respect to the two reference directions fixed by the vectors **A** and **C**. There will be a geodetic line through P for each of these ∞^1 directions, remembering that a geodetic line in a V_n is uniquely determined by one of its points and its direction at the point (cf. Art. 97). The aggregate of these geodetic lines constitutes what is called a Geodetic Surface of the V_n with Pole at P.

The Gaussian curvature of this geodetic surface is called the Riemannian Curvature of the V_n at P with respect to the section determined by the vectors **A** and **C**.

§103

Equations for an Infinitesimal Parallel Displacement of a Vector in a V_n

If a surface-vector **A** at a point $P(u^1, u^2)$ of a surface S undergo an infinitesimal parallel displacement to a point $P'(u^1 + du^1, u^2 + du^2)$, it becomes coincident with its equipollent vector at P', and in this displacement it will experience no absolute change in the sense defined in Art. 100. Hence, a condition for the infinitesimal parallel displacement of the surface vector **A** is that its absolute differential shall be zero; and if, conversely, this condition is satisfied in an infinitesimal displacement of a surface-vector **A**, then the displacement must be a parallel displacement.

An infinitesimal parallel displacement of a vector **A** at a point $P(u^1, u^2 \quad u^n)$ to a point $P'(u^1 + du^1, u^2 + du^2 \cdot \quad \cdot\; u^n + du^n)$ of a general V_n is defined by the single requirement: $d\mathbf{A} = 0$, where the symbol d indicates absolute differentiation.

For such displacement of the vector **A** from the point P to the point P' we must have, in accordance with equations (12) and (13), Art. 100:

$$dA^l = -A^i\{ik, l\}\, du^k, \tag{1}$$

$$dA_\rho = A_\lambda\{\rho\kappa, \lambda\}\, du^k. \tag{2}$$

But, from equations (6) and (7) of the same article, we find:

$$\{ik, l\}\, du^k = -\mathfrak{a}_i \cdot d\mathfrak{a}^l, \qquad \{\rho\kappa, \lambda\} = \mathfrak{a}^\lambda \cdot d\mathfrak{a}_\rho.$$

Hence:

$$dA^l = A^i \mathfrak{a}_i \cdot d\mathfrak{a}^l, \qquad dA_\rho = A_\lambda \mathfrak{a}^\lambda \cdot d\mathfrak{a}_\rho,$$

or:

(3) $$dA^l = \mathbf{A} \cdot d\mathfrak{a}^l;$$

(4) $$dA_\rho = \mathbf{A} \cdot d\mathfrak{a}\,.$$

The following equations, derivable from equations (1) and (2), for the partial derivatives of the contravariant and covariant components of the V_n vector $\mathbf{A}$, are valid in the case of the infinitesimal parallel displacement of $\mathbf{A}$:

(5) $$\frac{\partial A^l}{\partial u^k} = -\{ik, l\}\, A^i;$$

(6) $$\frac{\partial A_\rho}{\partial u^k} = \{\rho\kappa, \lambda\}\, A_\lambda.$$

The identifying indices in the equations of the present article are, of course, assumed to cover the integer range of values from 1 to n.

EXERCISES ON CHAPTER IX

1. Suppose given a quadratic form:

$$f = a_{ij}\, x_i\, x_j, \qquad (i, j = 1, 2 \ldots n),$$

which it is impossible to reduce by linear transformation to a form in less than n variables. The determinant of this form is:

$$D_n \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & & & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_n \end{vmatrix}$$

From this determinant the following series of determinants can be formed:

$$D_1 = a_{11},\ D_2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \text{ etc. to } D_n.$$

Show first that by a linear transformation f can be reduced to the form:

$$f = A_1 y_1^2 + A_2^2 y_2^2 + \cdots A_n y_n^2.$$

If in this expression the A-coefficients are all of the same sign, the form f is said to be **definite,** since it cannot change sign with any change of the variables. Show next, with the aid of a text on higher algebra, if necessary, that the form f will be **positive definite** provided:

$$D_k > 0, \qquad (k = 1, 2, \ldots n).$$

2. If in the case of a V_n the fundamental differential quadratic form can by a linear transformation be reduced to a sum of squares, the V_n is said to be Euclidean. Just as a non-developable surface can be regarded as embedded in a Euclidean V_3, show that any non-Euclidean V_n can be regarded as embedded in a Euclidean V_N, provided

$$N \geq \frac{1}{2}n(n+1).$$

3. In the case of any V_n it is always possible to find co-ordinates which are such that in the immediate vicinity of any assigned point they behave as rectangular Cartesian co-ordinates in the sense that the metrical coefficients become stationary; such co-ordinates are called Geodetic or Locally Cartesian co-ordinates. Verify the proposition here stated, considering first the special case of a non-Euclidean V_2.

4. On a cylindrical co-ordinate system in a Euclidean V_3, it is known that:

$$g_{11} = 1,\ g_{22} = \rho^2,\ g_{33} = 1, \qquad g_{ij} = 0 \text{ for } j \neq i.$$

Show that the non-vanishing Christoffel symbols of the second kind have the following values:

$$\{22, 1\} = -\rho,\ \{21, 2\} = \{12, 2\} = 1.$$

5. On a spherical co-ordinate system in a Euclidean V_3, it is known that:

$$g_{11} = 1,\ g_{22} = r^2 \sin^2\theta,\ g_{33} = r^2, \qquad g_{ij} = 0 \text{ for } j \neq i.$$

Show that the non-vanishing Christoffel 3-index symbols of the second kind have the following values:

$$\{22, 1\} = -r^2 \sin^2\theta,\ \{22, 3\} = -\sin\theta\cos\theta,\ \{33, 1\} = -r,$$

$$\{21, 2\} = \{12, 2\} = \frac{1}{r},\ \{31, 3\} = \{13, 3\} = \frac{1}{r},$$

$$\{32, 2\} = \{23, 2\} = \operatorname{ctn}\theta.$$

6. Show that the differentials of the metrical coefficients for any V_n satisfy the following n^2 equations:

$$dg_{ij} = [g_{\alpha j}\{i\beta, \alpha\} + g_{\alpha i}\{j\beta, \alpha\}]\,du^\beta, \quad (i, j, \alpha, \beta = 1, 2, \ldots n).$$

7. Show that the absolute differential of a V_n-idemfactor vanishes.

CHAPTER X

TENSOR THEORY

§104

Multilinear Forms

As was seen in Art. 92, a vector associated with any point P of a V_n can be represented as an invariant linear form in the base-vectors at P whose n coefficients are pure numbers which specify the numerical aspect of the vector. A generalization is now contemplated by which we shall be led to the definition of a tensor of rank r at a point P of a V_n as an invariant r-linear form in the base-vectors at P whose n^r scalar coefficients specify the numerical aspect of the tensor.

As usual, let $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_n$ denote the n unitary vectors on an arbitrary U-system of co-ordinates at a point $P(u^1, u^2 \ldots, u^n)$ in a V_n. As an example of a multi-linear form we take the trilinear form:[1)]

$$(1) \qquad A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k,$$

where the A-coefficients are scalar quantities whose values for the present can be considered as arbitrary; in this form the indeterminate or base products, of which $\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k$ is typical, are supposed subject to the associative but not to the commutative law. The form (1) involves unitary vectors only, but multi-linear forms which involve both unitary and reciprocal unitary vectors, or reciprocal unitary vectors only, will also be considered.

It will be assumed that in the addition of the terms of a multilinear form the commutative law is valid; that the multiplication of the form by a scalar is equivalent to multiplication of the coefficient or of any base-vector of each term by the scalar; and that in the multiplication of base-vectors the distributive law is valid.

1) In the present chapter it is to be understood, unless otherwise specified, that the summation law is operative; that all literal identifying indices cover the integer range from 1 to n inclusive; and that indices enclosed in parentheses have neither covariant nor contravariant significance.

With this understanding any unitary vector which appears in a multi-linear form can be replaced by the corresponding reciprocal unitary vector by the process of raising a subscript, and vice versa, as will be shown in detail in Art. 106.

The following statement constitutes a criterion of equality of two r-linear forms:

Two r-linear forms are equal if, when expressed on the same base-system, the coefficients of terms with the same base products are equal.

For example, if $\mathbf{A}_{(3)}$ and $\mathbf{B}_{(3)}$ denote two trilinear forms which can be expressed as follows:

$$\mathbf{A}_{(3)} = A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k, \qquad \mathbf{B}_{(3)} = B^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k,$$

then the two forms will be equal if, for all values of i, j, k:

$$A^{ijk} = B^{ijk}.$$

§105

Definition of a Tensor

A Tensor of Rank r associated with a point P of a V_n is an r-linear form in the base-vectors associated with the point whose coefficients are in general functions of the co-ordinates of the point and which is an invariant to choice of co-ordinate system.

As a consequence of this definition of a tensor it follows, as will be seen, that if a set of values are assigned to the coefficients of a tensor on any system of co-ordinates, then the coefficients of the tensor on any other system of co-ordinates can be calculated by a rule which is determined by the invariant condition imposed upon the tensor by its definition.

For the present it will suffice by way of elucidation of the definition of a tensor to discuss the special case of a tensor of the third rank. Consider, then, the trilinear form

$$A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k,$$

expressed on a U-system of co-ordinates, which becomes on a V-system

$$B^{lmn}\mathfrak{b}_l\mathfrak{b}_m\mathfrak{b}_n.$$

Then the trilinear form will represent a tensor of the third rank provided:

$$B^{lmn}\mathfrak{b}_l\mathfrak{b}_m\mathfrak{b}_n = A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k.$$

This requires that the A-coefficients of the form shall transform in accordance with definite rules. For, if the unitary vectors in the expression on the right be transformed in accordance with equations (II), Art. 95, the expression becomes a tri-linear form on the V-system, and the condition for equality of two multi-linear forms stated in the last paragraph of the preceding article requires that each B-coefficient of a unitary product on the left of the last equation shall be equal to the coefficient of the term on the right which involves the same unitary product, attention being paid, of course, to the order of the indices.[1]

The typical term of the tensor of the third rank under consideration, viz:

$$A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k, \qquad \text{(no summation)},$$

consists of the product of two parts, viz:

$$A^{ijk} \quad \text{and} \quad \mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k.$$

The first part is a scalar quantity which will be called the Component of the tensor;[2] the second part is an indeterminate product of a number of base-vectors equal to the rank of the tensor. Since, in virtue of its definition, a tensor is an invariant, each of its components must transform contragrediently with respect to each of the base-vectors of the indeterminate product with which it is associated; in the special case of the tensor of the third rank under consideration, the component A^{ijk} must be contravariant with respect to the indices i, j, k, a result which has been anticipated by the notation.

The number of terms in a tensor of the third rank is equal to n^3, this representing the number of permutations with repetition of n things taken 3 at a time, and the tensor will, of course, have the same number of components.

Similar statements to those made above relative to a tensor of rank 3 will, of course, apply to a tensor of rank r. The number of terms in a tensor of rank r is equal to n^r, and it will therefore have n^r components.

A tensor of the first rank is of particular interest in that it is nothing more nor less than a V_n vector.

A scalar point function in a V_n is an invariant, and can be classed as a tensor of zero rank.

[1] The general rules for the transformation of the coefficients of tensor forms will be found and discussed in Art. 107.

[2] The term "component" instead of "measure-number" is used in order to conform with common usage of writers on tensor analysis.

It is important to recognize the significance of the order in which the identifying indices appear in the components and the base-products of the terms of a tensor. If the order of the indices in the components of each term of a tensor be changed, a new tensor of the same rank will in general be obtained unless the same change of order of the indices in the corresponding base-product of each term be made. Thus, for example, in general:

$$A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k \neq A^{jki}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k;$$

on the other hand, by the dummy index rule:

$$A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k = A^{jki}\mathfrak{a}_j\mathfrak{a}_k\mathfrak{a}_i.$$

The tensor on the right of the inequality sign will, to be sure, have the same array of components as the tensor on the left, but the individual components of the two tensors will in general be associated with different base-products in the terms of the tensors, and, consequently, they will not in general be equal.

In virtue of its definition any tensor possesses certain attributes which find their origin in the fundamental differential quadratic form which determines the metrical properties of the V_n-manifold which, through the base-vectors, is involved in the definition of the tensor. This form for a general V_n is:

$$\overline{ds}^2 = g_{ij}du^i du^j, \qquad (g_{ij} = g_{ji}),$$

where:

(2) $$g_{ij} = \mathfrak{a}_i \cdot \mathfrak{a}_j.$$

Any attribute of a tensor which is dependent upon the relations expressed by equation (2) is a Metrical Property of the tensor, and any property of the tensor which does not depend upon these relations is a Non-Metrical Property.

Consider, for example, the simple case of a V_n-vector:

$$\mathbf{A} = A^i\mathfrak{a}_i = A^j\mathfrak{a}_j.$$

Upon taking the direct product of $\mathbf{A}$ into itself, we get:

$$\begin{aligned}\mathbf{A}\cdot\mathbf{A} &= A^2\\ &= \mathfrak{a}_i\cdot\mathfrak{a}_j A^i A^j,\end{aligned}$$

or, in virtue of equations (2):

(3) $$A^2 = g_{ij}A^iA^j.$$

This equation for the square of the magnitude of the vector expresses a metrical property of the vector.

On the other hand, the equations of transformation for the components of the vector, which hold in passing from one co-ordinate system to another, express non-metrical properties of the vector.

§106

Various Forms of a Tensor

It will be convenient to designate the form of a general tensor in which only unitary vectors appear as a Unitary Form; a form in which both unitary and reciprocal unitary vectors appear will be designated a Mixed Form; and a form in which only reciprocal unitary vectors appear will be designated a Reciprocal Form.

If a tensor is given in any one of these three forms, it can be changed into any other form with the aid of the relationships of unitary and reciprocal unitary vectors expressed by the familiar equations:

(a) $$\mathfrak{a}^\lambda = g^{i\lambda}\mathfrak{a}_i,$$ (b) $$\mathfrak{a}_i = g_{\lambda i}\mathfrak{a}^\lambda.$$

For example, let us start with a tensor $\mathbf{A}_{(3)}$ of the third rank expressed in the unitary form, viz:

(1) $$\mathbf{A}_{(3)} = A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k.$$

Evidently, with the aid of equations (b), we can write:

(2) $$\mathbf{A}_{(3)} = A_{\lambda}^{\cdot jk}\mathfrak{a}^\lambda\mathfrak{a}_j\mathfrak{a}_k,$$

where:

(3) $$A_{\lambda}^{\cdot jk} = g_{\lambda i}A^{ijk}.$$

We have thus expressed the tensor $\mathbf{A}_{(3)}$ in a mixed form. The process whereby the mixed component $A_{\lambda}^{\cdot jk}$ is derived from A^{ijk} is that of "lowering" an index of a component of the tensor. The process is reversible; for starting with the tensor in the mixed form (2), with the aid of equation (a), we get:

(4) $$\mathbf{A}_{(3)} = A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k,$$

provided:

(5) $$A^{ijk} = g^{i\lambda}A_{\lambda}^{\cdot jk}.$$

The process whereby A^{ijk} is derived from $A_{\lambda}^{\cdot jk}$ is that of "raising" an index of a component of the tensor. Equations (3) and (5) must, of course, be compatible, as can be shown as follows: multiplying equation (5) by g_{il}, we get:

$$g_{il}A^{ijk} = g_{il}g^{i\lambda}A_{\lambda}^{\cdot jk},$$

and since $g_{il}g^{i\lambda} = 1$ or 0, according as $l = \lambda$ or $l \neq \lambda$, it follows that:

$$A_{\lambda\cdot\cdot}^{\cdot jk} = g_{\lambda i}A^{ijk},$$

which is equation (3).

With the aid of the principles of raising and lowering indices the tensor $\mathbf{A}_{(3)}$ can be expressed in the following eight essentially different forms:

$$(6)\qquad \begin{aligned}\mathbf{A}_{(3)} &= A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k = A_{\lambda\cdot\cdot}^{\cdot jk}\mathfrak{a}^\lambda\mathfrak{a}_j\mathfrak{a}_k\\ &= A^{i\cdot k}_{\cdot\mu\cdot}\mathfrak{a}_i\mathfrak{a}^\mu\mathfrak{a}_k = A^{ij\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu\\ &= A_{\lambda\mu\cdot}^{\cdot\cdot k}\mathfrak{a}^\lambda\mathfrak{a}^\mu\mathfrak{a}_k = A^{i\cdot\cdot}_{\cdot\mu\nu}\mathfrak{a}_i\mathfrak{a}^\mu\mathfrak{a}^\nu\\ &= A_{\lambda\cdot\nu}^{\cdot j\cdot}\mathfrak{a}^\lambda\mathfrak{a}_j\mathfrak{a}^\nu = A_{\lambda\mu\nu}\mathfrak{a}^\lambda\mathfrak{a}^\mu\mathfrak{a}^\nu.\end{aligned}$$

Of these forms the first is purely unitary, and the last purely reciprocal; the other six are all mixed forms. The "dot" notation used in designating a component of a mixed form makes possible an unambiguous indication by the notation of the order in which the various indices appear. In the case of a unitary, or of a reciprocal form, the dot notation is, of course, not demanded.

The discussion given above for a tensor of the third rank can be easily extended so as to cover the case of a tensor of any rank.

§107

Transformation Equations for Tensor Components

Each term of a tensor consists of the product of one of its components and an indeterminate product of base-vectors, and, the laws of transformation for the base-vectors being known, those for the components can be derived in virtue of the invariance of the tensor itself. The process involved has already been indicated in Art. 95 in the derivation of the transformation equations for a V_n vector (tensor of the first rank).

Consider a U and a V-system of co-ordinates, the unitary vectors on the U-system being denoted as usual by $\mathfrak{a}_1, \mathfrak{a}_2, \ldots \mathfrak{a}_n$ and the reciprocal unitary vectors by $\mathfrak{a}^1, \mathfrak{a}^2 \ldots \mathfrak{a}^n$, while on the V-system the corresponding vectors are denoted by $\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_n$ and by $\mathbf{b}^1, \mathbf{b}^2, \ldots \mathbf{b}^n$ respectively.

Let $\mathbf{A}_{(1)}$, $\mathbf{A}_{(2)}$, $\mathbf{A}_{(3)}$ denote tensors of the first, second, and third rank, and let these tensors be expressed in the following various forms:

$$\begin{aligned}\mathbf{A}_{(1)} &= A^i\mathfrak{a}_i = A_\lambda\mathfrak{a}^\lambda;\\ \mathbf{A}_{(2)} &= A^{ij}\mathfrak{a}_i\mathfrak{a}_j = A_{\lambda\mu}\mathfrak{a}^\lambda\mathfrak{a}^\mu = A^{i\cdot}_{\cdot\mu}\mathfrak{a}_i\mathfrak{a}^\mu;\\ \mathbf{A}_{(3)} &= A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k = A_{\lambda\mu\nu}\mathfrak{a}^\lambda\mathfrak{a}^\mu\mathfrak{a}^\nu = A^{ij\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu.\end{aligned}$$

The components of these tensors will be denoted on the V-system by B's.

The transformation equations for the various components of the three tensors under consideration can be shown to be as follows:

(I) $$B^l = \frac{\partial v^l}{\partial u^i} A^i, \qquad A^i = \frac{\partial u^i}{\partial v^l} B^l;$$

(I′) $$B_\rho = \frac{\partial u^\lambda}{\partial v^\rho} A_\lambda, \qquad A_\lambda = \frac{\partial v^\rho}{\partial u^\lambda} B_\rho;$$

(II) $$B^{lm} = \frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j} A^{ij}, \qquad A^{ij} = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m} B^{lm};$$

(II′) $$B_{\rho\sigma} = \frac{\partial u^\lambda}{\partial v^\rho}\frac{\partial u^\mu}{\partial v^\sigma} A_{\lambda\mu}, \qquad A_{\lambda\mu} = \frac{\partial v^\rho}{\partial u^\lambda}\frac{\partial v^\sigma}{\partial u^\mu} B_{\rho\sigma};$$

(II″) $$B^{l\cdot}_{\cdot\sigma} = \frac{\partial v^l}{\partial u^i}\frac{\partial u^\mu}{\partial v^\sigma} A^{i\cdot}_{\cdot\mu}, \qquad A^{i\cdot}_{\cdot\mu} = \frac{\partial u^i}{\partial v^l}\frac{\partial v^\sigma}{\partial u^\mu} B^{l\cdot}_{\cdot\sigma};$$

(III) $$B^{lmn} = \frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j}\frac{\partial v^n}{\partial u^k} A^{ijk}, \qquad A^{ijk} = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m}\frac{\partial u^k}{\partial v^n} B^{lmn}$$

(III′) $$B_{\rho\sigma\tau} = \frac{\partial u^\lambda}{\partial v^\rho}\frac{\partial u^\mu}{\partial v^\sigma}\frac{\partial u^\nu}{\partial v^\tau} A_{\lambda\mu\nu}, \qquad A_{\lambda\mu\nu} = \frac{\partial v^\rho}{\partial u^\lambda}\frac{\partial v^\sigma}{\partial u^\mu}\frac{\partial v^\tau}{\partial u^\nu} B_{\rho\sigma\tau};$$

(III″) $$B^{lm\cdot}_{\cdot\cdot\tau} = \frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j}\frac{\partial u^\nu}{\partial v^\tau} A^{ij\cdot}_{\cdot\cdot\nu}, \qquad A^{ij\cdot}_{\cdot\cdot\nu} = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m}\frac{\partial v^\tau}{\partial u^\nu} B^{lm\cdot}_{\cdot\cdot\tau}$$

It will suffice, by way of exemplification of the methods of derivation of these transformation equations, to show that equations (II″) are valid: Since the tensor $\mathbf{A}_{(2)}$ is an invariant, we can write:

$$\begin{aligned} B^{l\cdot}_{\cdot\sigma}\mathbf{b}_l\mathbf{b}^\sigma &= A^{i\cdot}_{\cdot\mu}\mathbf{a}_i\mathbf{a}^\mu \\ &= A^{i\cdot}_{\cdot\mu}\frac{\partial v^l}{\partial u^i}\frac{\partial u^\mu}{\partial v^\sigma}\mathbf{b}_l\mathbf{b}^\sigma, \end{aligned}$$

with the aid of equations (II) and (IV), Art. 95. Hence, by the criterion of equality of two multi-linear forms:

$$B^{l\cdot}_{\cdot\sigma} = \frac{\partial v^l}{\partial u^i}\frac{\partial u^\mu}{\partial v^\sigma} A^{i\cdot}_{\cdot\mu}.$$

Similarly, the validity of the second of equations (II) can be shown.

Inspection of the above transformation equations for the components of tensors of the first, second, and third ranks will indicate how the appropriate transformation equations for the components, contravariant, covariant, or mixed, of a tensor of any rank can be written down by analogy.

An adequate criterion of a tensor form can be found through comparison of the transformation equations of the factors con-

stituting a typical term. For example, in the case of the tensor of the third rank:

$$\mathbf{A}_{(3)} = A^{ij\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu,$$

the transformation equations for the component of the typical term and for the corresponding indeterminate product are as follows:

$$B^{lm\cdot}_{\cdot\cdot\tau} = \frac{\partial v^l}{\partial u^i}\frac{\partial v^m}{\partial u^j}\frac{\partial u^\nu}{\partial v^\tau} A^{ij\cdot}_{\cdot\cdot\nu},$$

$$\mathbf{b}_l\mathbf{b}_m\mathbf{b}^\tau = \frac{\partial u^i}{\partial v^l}\frac{\partial u^j}{\partial v^m}\frac{\partial v^\tau}{\partial u^\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu.$$

These equations show that the component $A^{ij\cdot}_{\cdot\cdot\nu}$ transforms contragrediently with respect to the unitary vectors $\mathfrak{a}_i$, $\mathfrak{a}_j$, and the reciprocal unitary vector $\mathfrak{a}^\nu$.

In general, the component of the typical term of a tensor of the r'th rank has r indices, and to each index there corresponds a base-vector in the indeterminate product of the term with respect to which the component transforms contragrediently. This property is necessary and sufficient for the identification of a tensor form.

Since the equations of transformation for the components of a tensor are linear and homogeneous, it follows that if all the components of a tensor vanish on any system of co-ordinates they must vanish on all systems; this property can also be inferred directly from the definition of a tensor given in Art. 105. In the applications of tensor theory this property is of great importance; for example, if a physical law on a given system of co-ordinates is expressed by the statement that the components of a certain tensor vanish, then the same law on a new system of co-ordinates is expressed by stating that the components of the tensor on the new system vanish.

Symmetric and Anti-Symmetric Tensors

A tensor of the second or higher rank is said to be completely symmetric if the order of the indices in all of its components can be changed at will with no resulting change in the tensor itself.

The number of distinct components of a V_n tensor of rank r upon whose components no special conditions are imposed is n^r, this being the number of permutations with repetition of n things taken

r at a time. If conditions for complete symmetry are imposed upon the tensor, the number of its distinct components will be reduced from n^r to

$$\frac{n(n+1)(n+2)\cdots(n+r-1)}{r!},$$

since this expression represents the number of combinations with repetition of n things taken r at a time.

For example, consider the tensor $\mathbf{A}_{(2)}$ of the second rank, expressed in the unitary, mixed, and reciprocal forms:

$$\begin{aligned} \mathbf{A}_{(2)} &= A^{ij}\mathfrak{a}_i\mathfrak{a}_j, \\ &= A^{i\cdot}_{\cdot\mu}\mathfrak{a}_i\mathfrak{a}^\mu, \\ &= A_{\lambda\mu}\mathfrak{a}^\lambda\mathfrak{a}^\mu. \end{aligned} \tag{1}$$

This tensor will be symmetric provided that for all integer values of i, j, λ, and μ from 1 to n inclusive:

$$A^{ji} = A^{ij},$$

or:

$$A_{\mu\cdot}^{\cdot i} = A^{i\cdot}_{\cdot\mu},$$

or:

$$A_{\mu\lambda} = A_{\lambda\mu}.$$

A tensor of the second or higher rank is said to be completely anti-symmetric provided the interchange of any two indices in each component of the tensor changes its sign, but not otherwise.

A completely anti-symmetric tensor cannot, in virtue of its definition, contain any component in which two of the indices are the same. On this account the number of distinct components in a completely anti-symmetric tensor is less than the number of distinct components in a completely symmetric tensor of the same rank; it is in fact equal to

$$\frac{n(n-1)(n-2)\cdots(n-r+1)}{r!},$$

since this expression represents the number of combinations without repetition of n things taken r at a time.

The tensor $\mathbf{A}_{(2)}$ of the second rank, expressed in unitary, mixed, and reciprocal forms by equations (1), will be anti-symmetric provided:

$$A^{ji} = -A^{ij},$$

or:

$$A_{\mu\cdot}^{\cdot i} = -A^{i\cdot}_{\cdot\mu},$$

or:

$$A_{\mu\lambda} = -A_{\lambda\mu},$$

where the indices i, j, λ, and μ can assume any integer values from 1 to n inclusive.

A tensor of the third or higher rank may, of course, be symmetric or anti-symmetric with respect to certain pairs of indices, though not to all.

§109

Physical Example of a Tensor

Before proceeding further with the development of the theory of tensors we shall consider a physical example of a tensor of the second rank which may serve to throw further light on the real significance of the tensor idea.

The tensor which we shall consider makes its appearance in a discussion of the stress at a point of an elastic medium in a state of strain.[1] Referring to Fig. 42, $\underline{\mathbf{n}}$ represents a unit vector $\mathbf{n}$ whose direction is normal to an infinitesimal element of area ω of a surface S passing through the point $P(u^1, u^2, u^3)$ of an elastic medium, and which divides the medium into two portions, and $\underline{\mathbf{F}}$ represents the resultant $\mathbf{F}$ of the elastic forces exerted across the element of area by the portion of the medium toward which $\mathbf{n}$ is directed. The resultant stress across S at P, denoted by $\mathbf{R}$, is defined by the equation:

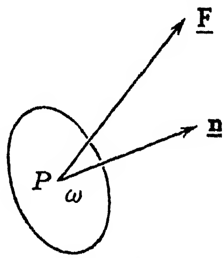


Fig. 42.

$$\mathbf{R} = \underset{\omega \to 0}{Lt} \frac{\mathbf{F}}{\omega}.$$

If, as we suppose, the elastic medium is in a state of equilibrium, there are two conditions which must be satisfied, the first of which states that the resultant of all the forces acting upon any element of volume must vanish, and the second that the resultant moment of these forces about any point must vanish.

For the element of volume we take an infinitesimal tetrahedron with one vertex at P and conterminous edges determined in direction by the unitary vectors associated with P, as shown in Fig. 43. The areas of the three faces to the tetrahedron which intersect in the point P are respectively designated by $dS_{(1)}$, $dS_{(2)}$, $dS_{(3)}$, the fourth face will be designated by dS, and a unit vector in the outward normal direction to this face will be denoted by $\mathbf{n}$. The

[1] It was in this connection that the term "tensor" was first used.

element of volume thus defined will be called the Unitary Tetrahedron.

Since the unitary tetrahedron is supposed infinitesimal, the forces acting upon it which are proportional to its volume can be neglected in comparison with those (the elastic forces) which are proportional to the areas of its faces. Hence, if $-\mathbf{R}_{(1)}$, $-\mathbf{R}_{(2)}$,

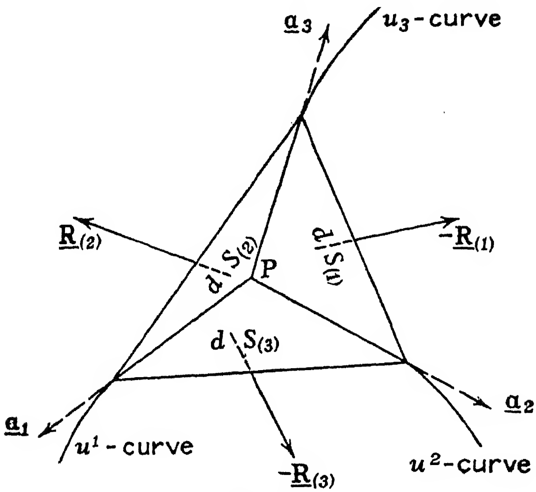


Fig. 43.

$-\mathbf{R}_{(3)}$, $\mathbf{R}$ represent the stresses in the outward direction across the faces of the unitary tetrahedron whose areas respectively are $dS_{(1)}$, $dS_{(2)}$, $dS_{(3)}$, dS, then, by the first condition for equilibrium, we shall have:[1)]

$$\mathbf{R}\, dS - \mathbf{R}_{(j)} dS_{(j)} = 0,$$

Now, by simple geometry:

$$dS_{(j)} = a^j \mathfrak{a}_j \cdot \mathbf{n} dS, \qquad \text{(no summation)},$$

where a^j denotes the magnitude of the reciprocal unitary vector $\mathfrak{a}^j$. We now define a new vector $\mathfrak{a}_{(j)}$ by the equation:

$$\mathfrak{a}_{(j)} = a^j \mathfrak{a}_j, \qquad \text{(no summation)}.$$

We can then write:

$$dS_{(j)} = \mathfrak{a}_{(j)} \cdot \mathbf{n} dS;$$

$$\mathbf{R} - \mathbf{R}_{(j)} \mathfrak{a}_{(j)} \cdot \mathbf{n} = 0.$$

[1)] Throughout the present article the summation convention is supposed operative unless otherwise stated, and identifying literal indices cover the integer range 1, 2, 3.

For the resultant stress $\mathbf{R}$ in the outward direction across the face dS of the unitary tetrahedron normal to the unit vector $\mathbf{n}$ we can then write:

$$\mathbf{R} = \mathbf{\Phi} \cdot \mathbf{n}, \tag{1}$$

where:

$$\mathbf{\Phi} = \mathbf{R}_{(j)}\mathbf{a}_{(j)}. \tag{2}$$

The dyadic $\mathbf{\Phi}$, which is here expressed in a trinomial form, is called the Stress Dyadic or Stress Tensor. It can easily be expressed in a form which shows at once its tensor character. To do this, we first express the vector $\mathbf{R}_{(j)}$ in terms of its components on the unitary base-system at P by writing:

$$\mathbf{R}_{(j)} = R^{\,i}_{(j)}\mathbf{a}_i,$$

where the coefficients of the unitary vectors on the right are, of course, the contravariant scalar components of $\mathbf{R}_{(j)}$. We can then write:

$$\mathbf{\Phi} = R^{\,i}_{(j)}\mathbf{a}_i\mathbf{a}_{(j)} = R^{\,i}_{(j)}a^j\mathbf{a}_i\mathbf{a}_j. \tag{3}$$

Next, we define a new scalar quantity S^{ij} by the equation:

$$S^{ij} = R^{\,i}_{(j)}a^j, \qquad \text{(not summed)}. \tag{4}$$

Then $\mathbf{\Phi}$ can be expressed in the unitary tensor form:

$$\mathbf{\Phi} = S^{ij}\mathbf{a}_i\mathbf{a}_j. \tag{5}$$

The nine S-coefficients, of which S^{ij} is typical, are contravariant with respect to both of the indices i and j, and will be called Contravariant Stress Parameters. When these nine stress parameters are known, the stress tensor and, consequently, by equation (1), the stress itself in the neighborhood of the point P will be completely determined. If it were only required to express the resultant stress across a single given plane at P, the number of parameters required would be only three, and the reason why more than three parameters are required for the complete specification of the stress at P is to be found in the fact that for such a specification a sufficient number of data must be given to permit of the calculation of the stress across every plane at P; for this purpose, as we have seen, nine data are sufficient, but among these, however, there are three relations, as will be seen.

So far no use has been made of the second condition of equilibrium. This condition requires that the resultant of the moments about any point of all the forces acting upon the unitary tetrahedron

shall vanish. In virtue of this condition it can easily be shown that the stress tensor must be symmetric, and hence that:

(6) $$S^{ij} = S^{ji}.$$

There are therefore three independent relations among the nine parameters which specify the stress tensor.

By interchanging the rôles of the unitary and the reciprocal unitary vectors, introducing as the element of volume the Reciprocal

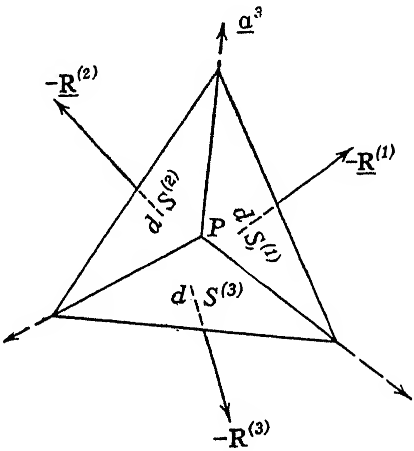


Fig. 44.

Tetrahedron shown in Fig. 44, and adopting a procedure quite analogous to that followed above, the following forms for $\mathbf{\Phi}$, analogous to the forms (2), (3), and (5), can be derived:

(2′) $$\mathbf{\Phi} = \mathbf{R}^{(\mu)}\mathfrak{a}^{(\mu)}, \qquad (\mathfrak{a}^{(\mu)} = a_\mu \mathfrak{a}^\mu \text{ not summed});$$

(3′) $$\mathbf{\Phi} = R^{(\mu)}_{(\lambda)}\mathfrak{a}^\lambda \mathfrak{a}^{(\mu)};$$

(5′) $$\mathbf{\Phi} = S_{\lambda\mu}\mathfrak{a}^\lambda \mathfrak{a}^\mu, \qquad (S_{\lambda\mu} = a_\mu R^{(\mu)}_{(\lambda)} \text{ not summed}).$$

The significance of the notation in these equations will be evident upon referring to Fig. 44, and to the analogous equations (2), (3), and (5).

Equation (5′) expresses the stress tensor in reciprocal form, in which the nine S-coefficients, of which $S_{\lambda\mu}$ is typical, are covariant with respect to both of the indices λ and μ. These coefficients are called Covariant Stress Parameters. Among them there exist the three independent relations:

(7) $$S_{\lambda\mu} = S_{\mu\lambda}.$$

These equations are analogous to equations (6) above, expressing the three independent relations among the nine contravariant stress parameters.

In case ordinary rectangular Cartesian co-ordinates are used, the system of reciprocal unitary vectors becomes identical with that of the unitary vectors themselves, and the distinctions between contravariant and covariant stress parameters disappear. In this case we can take:

$$\begin{array}{lll} \mathfrak{a}^1 = \mathfrak{a}_1 = \mathbf{i}, & \mathfrak{a}^2 = \mathfrak{a}_2 = \mathbf{j}, & \mathfrak{a}^3 = \mathfrak{a}_3 = \mathbf{k}, \\ S^{11} = S_{11} = P, & S^{22} = S_{22} = Q, & S^{33} = S_{33} = R, \\ S^{12} = S_{12} = U, & S^{23} = S_{23} = S, & S^{31} = S_{31} = T. \end{array}$$

The stress tensor then takes the form:

$$\begin{aligned} \mathbf{\Phi} = \; & P\mathbf{ii} + U\mathbf{ij} + T\mathbf{ik} \\ & + U\mathbf{ji} + Q\mathbf{jj} + S\mathbf{jk} \\ & + T\mathbf{ki} + S\mathbf{kj} + R\mathbf{kk}. \end{aligned}$$

The resultant stress $\mathbf{R}(=\mathbf{\Phi}\cdot\mathbf{n})$ can now be expressed in the form appropriate to rectangular Cartesian co-ordinates:

$$\begin{aligned} \mathbf{R} = \; & (P\mathbf{i}\cdot\mathbf{n} + U\mathbf{j}\cdot\mathbf{n} + T\mathbf{k}\cdot\mathbf{n})\mathbf{i} \\ & + (U\mathbf{i}\cdot\mathbf{n} + Q\mathbf{j}\cdot\mathbf{n} + S\mathbf{k}\cdot\mathbf{n})\mathbf{j} \\ & + (T\mathbf{i}\cdot\mathbf{n} + S\mathbf{j}\cdot\mathbf{n} + R\mathbf{k}\cdot\mathbf{n})\mathbf{k}. \end{aligned}$$

§110

Addition of Tensors of the Same Rank

The sum of two tensors of the same rank is equivalent to a new tensor of like rank.

To exemplify the addition of two tensors of the same rank we consider the special case of the addition of two tensors $\mathbf{A}_{(3)}$ and $\mathbf{B}_{(3)}$ of the third rank, which we suppose given in the non-identical forms:

$$\begin{array}{ll} \mathbf{A}_{(3)} & A^{ijk}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_k, \\ \mathbf{B}_{(3)} & B^{ij\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu. \end{array}$$

Before the process of addition can be carried out one or the other or both of these forms must be so modified as to obtain forms for $\mathbf{A}_{(3)}$ and $\mathbf{B}_{(3)}$ which are identical. This can be accomplished, for example, by rewriting $\mathbf{A}_{(3)}$ in the form:

$$\mathbf{A}_{(3)} = A^{ij\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu,$$

where:

$$A^{ij\cdot}_{\cdot\cdot\nu} = g_{\nu k}A^{ijk}.$$

For the sum of $\mathbf{A}_{(3)}$ and $\mathbf{B}_{(3)}$ we can then write:

$$\mathbf{C}_{(3)} = C^{ij\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}^\nu,$$

where:

$$C^{ij\cdot}_{\cdot\cdot\nu} = A^{ij\cdot}_{\cdot\cdot\nu} + B^{ij\cdot}_{\cdot\cdot\nu}.$$

Since $\mathbf{A}_{(3)}$ and $\mathbf{B}_{(3)}$ are both tensors and therefore invariants, it follows that $\mathbf{C}_{(3)}$ is an invariant and therefore a tensor; it is obviously of the same rank as $\mathbf{A}_{(3)}$ and $\mathbf{B}_{(3)}$.

Any two tensors of the same rank can be added in like manner.

§111
Multiplication of Tensors

In forming products of tensors the associative and distributive laws are assumed to be valid, provided the order of their base-vectors is maintained.

The consideration of a typical example will suffice to make clear the nature of the general process of forming a general product of any two tensors.

Consider the two tensors $\mathbf{A}_{(2)}$ and $\mathbf{B}_{(3)}$, of rank 2 and rank 3 respectively, expressed in the following forms:

(1) $$\mathbf{A}_{(2)} = A^{ij}\mathfrak{a}_i\mathfrak{a}_j,$$

(2) $$\mathbf{B}_{(3)} = B^{lm\cdot}_{\nu\cdot\cdot}\mathfrak{a}_l\mathfrak{a}_m\mathfrak{a}^\nu.$$

The general products of $\mathbf{A}_{(2)}$ into $\mathbf{B}_{(3)}$ and of $\mathbf{B}_{(3)}$ into $\mathbf{A}_{(2)}$ are, respectively, formed as follows:

(3) $$\mathbf{A}_{(2)}\mathbf{B}_{(3)} = A^{ij}B^{lm\cdot}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_l\mathfrak{a}_m\mathfrak{a}^\nu,$$

(4) $$\mathbf{B}_{(3)}\mathbf{A}_{(2)} = B^{lm\cdot}_{\cdot\cdot\nu}A^{ij}\mathfrak{a}_l\mathfrak{a}_m\mathfrak{a}^\nu\mathfrak{a}_i\mathfrak{a}_j.$$

The scalar coefficients in the expressions for each of the products are contravariant with respect to the indices i, j, l, m, and covariant with respect to the index ν, while the corresponding base-products are covariant with respect to the indices i, j, l, m, and contravariant with respect to the index ν. We conclude, therefore, that $\mathbf{A}_{(2)}\mathbf{B}_{(3)}$ and $\mathbf{B}_{(3)}\mathbf{A}_{(2)}$ are invariants and therefore tensors, each, evidently, of rank 5.

It should be specially noticed that the difference between the products $\mathbf{A}_{(2)}\mathbf{B}_{(3)}$ and $\mathbf{B}_{(3)}\mathbf{A}_{(2)}$ arises from the reverse order of appearance of the base products of the factors $\mathbf{A}_{(2)}$ and $\mathbf{B}_{(3)}$.

A different type of product of two tensors, called a Direct Product or Inner Product, is formed in the same manner as the general product except that in forming the product a dot is inserted

between the two indeterminate products of the two tensors, which signifies that the direct product of the two base-vectors between which the dot stands is to be formed and considered as a scalar factor of the rest of the expression. The direct product of two tensors of rank r and s respectively is itself a tensor of rank $r + s - 2$.

The consideration of a typical example will suffice to make clear the nature of the general process of forming a direct product of any two tensors.

Consider the two tensors $\mathbf{A}_{(2)}$ and $\mathbf{B}_{(3)}$ expressed in the forms (1) and (2). The direct product of $\mathbf{A}_{(2)}$ into $\mathbf{B}_{(3)}$ and of $\mathbf{B}_{(3)}$ into $\mathbf{A}_{(2)}$, denoted by $\mathbf{A}_{(2)} \cdot \mathbf{B}_{(3)}$ and $\mathbf{B}_{(3)} \cdot \mathbf{A}_{(2)}$ respectively, are then formed as follows:

$$\mathbf{A}_{(2)} \cdot \mathbf{B}_{(3)} = A^{ij} B^{lm\cdot}_{\cdot\cdot\nu} \mathfrak{a}_i \mathfrak{a}_j \cdot \mathfrak{a}_l \mathfrak{a}_m \mathfrak{a}^\nu = g_{jl} A^{ij} B^{lm\cdot}_{\cdot\cdot\nu} \mathfrak{a}_i \mathfrak{a}_m \mathfrak{a}^\nu, \tag{5}$$

$$\mathbf{B}_{(3)} \cdot \mathbf{A}_{(2)} = B^{lm\cdot}_{\cdot\cdot\nu} A^{ij} \mathfrak{a}_l \mathfrak{a}_m \mathfrak{a}^\nu \cdot \mathfrak{a}_i \mathfrak{a}_j = B^{lm\cdot}_{\cdot\cdot\nu} A^{\nu j} \mathfrak{a}_l \mathfrak{a}_m \mathfrak{a}_j. \tag{6}$$

That both of the direct products thus formed are tensors is easily verified by observing that their covariant and contravariant dimensions balance, showing that they are invariants; evidently, each product represents a tensor of rank $2 + 3 - 2$.

In general, of course, the direct product of two tensors is not a scalar quantity; but if both tensors are of rank 1, that is if they are both vectors, their direct product will be a scalar invariant which can be considered a tensor of zero rank.

§112

Contraction of Tensors

Consider a tensor of the third rank expressed in the mixed form:

$$\mathbf{A}_{(3)} = A^{ij\cdot}_{\cdot\cdot\nu} \mathfrak{a}_i \mathfrak{a}_j \mathfrak{a}^\nu.$$

From this tensor construct three new tensors:

$$\begin{aligned} \mathbf{A}_{(1)} &= A^{ij\cdot}_{\cdot\cdot\nu} \mathfrak{a}_j \cdot \mathfrak{a}^\nu \mathfrak{a}_i, & \mathbf{B}_{(1)} &= A^{ij\cdot}_{\cdot\cdot\nu} \mathfrak{a}^\nu \cdot \mathfrak{a}_i \mathfrak{a}_j, & \mathbf{C}_{(1)} &= A^{ij\cdot}_{\cdot\cdot} \mathfrak{a}_i \cdot \mathfrak{a}_j \mathfrak{a}^\nu, \\ &= A^{ij\cdot}_{\cdot\cdot j} \mathfrak{a}_i, & &= A^{ij\cdot}_{\cdot\cdot i} \mathfrak{a}_j, & &= A^{ij\cdot}_{\cdot\cdot\nu} g_{ij} \mathfrak{a}^\nu, \end{aligned} \tag{1}$$

by replacing the indeterminate product of the original tensor by a scalar product of two of its vectors multiplied into the third. That the quantities represented by $\mathbf{A}_{(1)}$, $\mathbf{B}_{(1)}$, and $\mathbf{C}_{(1)}$ are actually tensors is evident upon inspection of their covariant and contravariant dimensions. Each of the tensors thus formed is of rank 1, that is, of a rank lower by 2 than the rank of the original tensor.

In a similar manner we can form from a tensor of rank $r(\geq 2)$ other tensors of rank $r - 2$, by replacing the indeterminate prod-

uct of the tensor by the scalar product of any two of its vectors multiplied into the factor obtained by elimination of these two vectors.

This process, whereby from a tensor of a given rank other tensors of rank lower by 2 are derived, is known as that of Contraction of a Tensor.

The contraction of a tensor of the second rank yields a scalar invariant, that is, a tensor of zero rank.

Consider, for example, the tensor:

$$\mathbf{A}_{(2)} = A^{ij}\mathfrak{a}_i\mathfrak{a}_j.$$

Upon contraction we obtain:

$$\mathbf{A}_{(0)} = A^{ij}\mathfrak{a}_i \cdot \mathfrak{a}_j = A^{ij}g_{ij} = \text{A Scalar Invariant.} \tag{2}$$

It is worthy of note that the direct product of any two tensors can be regarded as a tensor which is the result of contraction of a tensor of higher rank by 2. For example, in the direct product of $\mathbf{A}_{(2)}$ into $\mathbf{B}_{(3)}$, given by equation (5), Art. 111, the product can be regarded as a contraction of the following tensor of the fifth rank:

$$\mathbf{C}_{(5)} = A^{ij}B^{lm}_{\cdot\cdot\nu}\mathfrak{a}_i\mathfrak{a}_j\mathfrak{a}_l\mathfrak{a}_m\mathfrak{a}^\nu.$$

The process of contraction of a tensor of given rank r (≥ 2) can be repeated until, if r is even, a tensor of zero rank is obtained, or until, if r is odd, a tensor of rank 1 is obtained.

§113

The Fundamental Tensor

The quantity $\mathbf{I}$ defined as follows:

$$\mathbf{I} = g_{\alpha\beta}\mathfrak{a}^\alpha\mathfrak{a}^\beta,$$

where the g-coefficients are the coefficients in the fundamental differential quadratic form for a V_n, is a tensor, as can be verified by writing down its transformation equations. This tensor is called the Fundamental Tensor for a V_n. We shall see that the tensor $\mathbf{I}$ is, in fact, nothing more nor less than a V_n dyadic idemfactor.

With the aid of equations (2), Art. 94, $\mathbf{I}$ can be expressed in reciprocal, mixed, and unitary forms as follows:

$$\mathbf{I} = g_{\alpha\beta}\mathfrak{a}^\alpha\mathfrak{a}^\beta, \tag{1}$$
$$\mathbf{I} = \mathfrak{a}_\alpha\mathfrak{a}^\alpha, \tag{2}$$
$$\mathbf{I} = \mathfrak{a}^\alpha\mathfrak{a}_\alpha, \tag{3}$$
$$\mathbf{I} = g^{\alpha\beta}\mathfrak{a}_\alpha\mathfrak{a}_\beta. \tag{4}$$

In the form (1) the g-coefficients are doubly covariant components of **I**, and in the form (4) they are doubly contravariant components.

The direct product of **I** into any tensor is equal to the tensor itself. For example, forming the direct product of **I** into a tensor $\mathbf{A}_{(3)}$ of rank 3 expressed in unitary form, using the mixed form (2) for **I**, we find:

$$
\begin{aligned}
\mathbf{I} \cdot \mathbf{A}_{(3)} &= \mathfrak{a}_\alpha \mathfrak{a}^\alpha \cdot (A^{ijk} \mathfrak{a}_i \mathfrak{a}_j \mathfrak{a}_k) \\
&= A^{ijk} \mathfrak{a}_\alpha \mathfrak{a}^\alpha \cdot \mathfrak{a}_i \mathfrak{a}_j \mathfrak{a}_k \\
&= A^{ijk} \mathfrak{a}_i \mathfrak{a}_j \mathfrak{a}_k \\
&= \mathbf{A}_{(3)}.
\end{aligned} \tag{5}
$$

In like manner the direct product of **I**, used as prefactor or postfactor, and any tensor can be shown to be equal to the tensor itself.

§114
Differential V_n-Invariants

We define a differentiating operator ∇ for any point of a V_n, which for the special case of a Euclidean V_3 is equivalent to the operator ∇ discussed in Art. 88, as follows:

$$
\nabla = \mathfrak{a}^\alpha \frac{\partial}{\partial u^\alpha}, \tag{1}
$$

where the differentiating symbol indicates absolute differentiation.

This operator is itself a differential invariant. For, by the second of equations (IV), Art. 95:

$$
\mathfrak{a}^\alpha = \mathfrak{b}^\epsilon \frac{\partial u^\alpha}{\partial v^\epsilon};
$$

furthermore:

$$
\frac{\partial}{\partial u^\alpha} = \frac{\partial v^\beta}{\partial u^\alpha} \frac{\partial}{\partial v^\beta};
$$

so that:

$$
\mathfrak{a}^\alpha \frac{\partial}{\partial u^\alpha} = \mathfrak{b}^\epsilon \frac{\partial u^\alpha}{\partial v^\epsilon} \frac{\partial v^\beta}{\partial u^\alpha} \frac{\partial}{\partial v^\beta};
$$

and hence, in virtue of equation (4), Art. 95:

$$
\mathfrak{a}^\alpha \frac{\partial}{\partial u^\alpha} = \mathfrak{b}^\beta \frac{\partial}{\partial v^\beta}. \tag{2}
$$

It should be noticed that in the defining equation (1) for the operator ∇ the index α is a dummy index; it can, therefore, be replaced by any other index if so desired.

By operating with the invariant ∇ upon scalar and tensor point functions of a V_n, which are themselves invariants, various important differential invariants can be obtained. In operating with ∇ upon tensor point functions, it is to be understood that the operational laws of ordinary differentiation are followed, but with the proviso that in the differentiation of products of base vectors of a tensor function the order of appearance of the base vectors is not altered.

If U and V denote two scalar point functions, we find:

$$\nabla U = \mathfrak{a}^\lambda \frac{\partial U}{\partial u^\lambda}, \qquad \nabla V = \mathfrak{a}^\mu \frac{\partial V}{\partial u^\mu}, \tag{3}$$

representing the Gradients of these functions.

Forming the direct products of these two invariants, we obtain another differential invariant:

$$\nabla U \cdot \nabla V = g^{\lambda\mu} \frac{\partial U}{\partial u^\lambda} \frac{\partial V}{\partial u^\mu}. \tag{4}$$

This is known as a Mixed Beltrami Differential Parameter.

In the special case for which $U = V$ we have:

$$\nabla U \cdot \nabla U = g^{\lambda\mu} \frac{\partial U}{\partial u^\lambda} \frac{\partial U}{\partial u^\mu} \tag{5}$$

This is called the First Beltrami Differential Parameter of U.

Consider now a V_n-vector (tensor of the first rank) expressed in the unitary form:

$$\mathbf{A} = A^i \mathfrak{a}_i.$$

The result of operating upon this vector with ∇ is to produce the tensor:

$$\nabla \mathbf{A} = \mathfrak{a}^\rho \left[\frac{\partial A^i}{\partial u^\rho} \mathfrak{a}_i + A^k \frac{\partial \mathfrak{a}_k}{\partial u^\rho} \right];$$

or, with the aid of equation (8), Art. 100:

$$\nabla \mathbf{A} = \left[\frac{\partial A^i}{\partial u^\rho} + \{\rho k, i\}\, A^k \right] \mathfrak{a}^\rho \mathfrak{a}_i.$$

Upon contraction of this tensor of the second rank we obtain the important scalar differential invariant:

$$\nabla \cdot \mathbf{A} = \frac{\partial A^i}{\partial u^i} + \{ik, i\}\, A^k. \tag{5}$$

This invariant is called the Divergence of the vector $\mathbf{A}$. It can easily be expressed in a form which justifies this designation as

follows: By the dummy index rule and by equation (8), Art. 98, we find that:

$$\{ik, i\}\, A^k = \{ki, k\}\, A^i = \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}}{\partial u^i}A^i;$$

and equation (5) can therefore be written in the form:

$$\nabla\cdot\mathbf{A} = \frac{1}{\sqrt{g}}\frac{\partial}{\partial u^i}(\sqrt{g}A^i). \tag{6}$$

This expression reduces for a Euclidean V_3 to that found for the divergence of a vector in general co-ordinates and expressed by equation (2*c*), Art. 88.

If the vector $\mathbf{A}$ be expressed in the reciprocal form:

$$\mathbf{A} = A_\lambda \mathbf{a}^\lambda,$$

then, with the aid of equation (9), Art. 100, we find the following alternative form for the divergence of $\mathbf{A}$:

$$\nabla\cdot\mathbf{A} = g^{\rho\lambda}\left[\frac{\partial A_\lambda}{\partial u^\rho} - \{\rho\lambda, \epsilon\}\, A_\epsilon\right]. \tag{7}$$

In particular, if $\mathbf{A}$ is the gradient of a scalar point function U, so that:

$$\mathbf{A} = \nabla U, \qquad \vdots\ \frac{\partial U}{\partial u^\lambda}, \tag{8}$$

then, by the preceding formula:

$$\nabla\cdot\nabla U = \Delta U = g^{\lambda\rho}\left[\frac{\partial^2 U}{\partial u^\lambda \partial u^\rho} - \{\lambda\rho, \epsilon\}\frac{\partial U}{\partial u^\epsilon}\right]. \tag{9}$$

This is known as Lame's or the Second Beltrami Differential Parameter of U. With the aid of equations (7) and (8), Art. 98, it can easily be shown to be capable of expression in the more concise form:

$$\nabla\cdot\nabla U = \Delta U = \frac{1}{\sqrt{g}}\frac{\partial}{\partial u^\rho}\left(\sqrt{g}\, g^{\lambda\rho}\frac{\partial U}{\partial u^\lambda}\right). \tag{10}$$

§115

Covariant Differentiation of Tensor Components

The operator ∇ acting upon any tensor produces a new tensor of the next higher rank. Any component of the new tensor thus obtained is called the Covariant Derivative of a corresponding component of the original tensor. The method of covariant differentiation of tensor components will be sufficiently exemplified by

showing in detail how the covariant derivatives of the covariant components of tensors of the first and second rank are obtained.

Consider the vector $\mathbf{A}_{(1)}$ (tensor of the first rank):

$$\mathbf{A}_{(1)} = A_\mu \mathfrak{a}^\mu.$$

Operating with ∇ upon $\mathbf{A}_{(1)}$ we obtain a tensor $\mathbf{A}_{(2)}$ of the second rank as follows:

$$\begin{aligned}\mathbf{A}_{(2)} = \nabla \mathbf{A}_{(1)} &= \mathfrak{a}^\rho \frac{\partial}{\partial u^\rho}(A_\mu \mathfrak{a}^\mu) \\ &= \mathfrak{a}^\rho \left[\frac{\partial A_\mu}{\partial u^\rho}\mathfrak{a}^\mu + A_\mu \frac{\partial \mathfrak{a}^\mu}{\partial u^\rho}\right] \\ &= \mathfrak{a}^\rho \left[\frac{\partial A^\mu}{\partial u^\rho}\mathfrak{a}^\mu - A_\mu \{\epsilon\rho, \mu\} \mathfrak{a}^\epsilon\right],\end{aligned}$$

with the aid of equation (9), Art. 100; and upon interchanging the dummy indices ϵ and μ in the last term of the third line we can write:

$$\mathbf{A}_{(2)} = \nabla \mathbf{A}_{(1)} = A_{\mu\rho}\mathfrak{a}^\rho \mathfrak{a}^\mu,$$

where:

$$A_{\mu\rho} = \frac{\partial A_\mu}{\partial u^\rho} - \{\mu\rho, \epsilon\} A_\epsilon.$$

This doubly covariant component of the tensor $\mathbf{A}_{(2)}$ is the covariant derivative of the covariant component A_μ of the vector $\mathbf{A}_{(1)}$.

Consider next the tensor of the second rank:

$$\mathbf{B}_{(2)} = B_{\mu\rho}\mathfrak{a}^\rho \mathfrak{a}^\mu.$$

Operating with ∇ upon this tensor we obtain a new tensor $\mathbf{B}_{(3)}$ of the third rank as follows:

$$\begin{aligned}\mathbf{B}_{(3)} = \nabla \mathbf{B}_{(2)} &= \mathfrak{a}^\sigma \frac{\partial}{\partial u^\sigma}(B_{\mu\rho}\mathfrak{a}^\rho \mathfrak{a}^\mu) \\ &= \mathfrak{a}^\sigma \left[\frac{\partial B_{\mu\rho}}{\partial u^\sigma}\mathfrak{a}^\rho \mathfrak{a}^\mu + B_{\mu\rho}\frac{\partial \mathfrak{a}^\rho}{\partial u^\sigma}\mathfrak{a}^\mu + B_{\mu\rho}\mathfrak{a}^\rho \frac{\partial \mathfrak{a}^\mu}{\partial u^\sigma}\right] \\ &= \mathfrak{a}^\sigma \left[\frac{\partial B_{\mu\rho}}{\partial u^\sigma}\mathfrak{a}^\rho \mathfrak{a}^\mu - B_{\mu\rho}\{\epsilon\sigma, \rho\}\mathfrak{a}^\epsilon \mathfrak{a}^\mu - B_{\mu\rho}\{\epsilon\sigma, \mu\}\mathfrak{a}^\rho \mathfrak{a}^\epsilon\right],\end{aligned}$$

with the aid of equation (9), Art. 100; and upon interchanging in the last line the dummy indices ϵ and ρ in the second term and ϵ and μ in the third term we can write:

$$\mathbf{B}_{(3)} = B_{\mu\rho\sigma}\mathfrak{a}^\sigma \mathfrak{a}^\rho \mathfrak{a}^\mu,$$

where:

$$B_{\mu\rho\sigma} = \frac{\partial B_{\mu\rho}}{\partial u^\sigma} - \{\mu\sigma, \epsilon\} B_{\epsilon\rho} - \{\rho\sigma, \epsilon\} B_{\mu\epsilon}.$$

This component of the tensor $\mathbf{B}_{(3)}$ is the covariant derivative of the covariant component $B_{\mu\rho}$ of the tensor $\mathbf{B}_{(2)}$.

The methods used above in finding the covariant derivatives of the covariant components of the tensors $\mathbf{A}_{(1)}$ and $\mathbf{B}_{(2)}$ of the first and second ranks can be used to find the covariant derivatives of their contravariant components, and also of the mixed components of the tensor $\mathbf{B}_{(2)}$.

Accordingly, if we have given a vector $\mathbf{A}_{(1)}$ expressed in the forms:

$$\mathbf{A}_{(1)} = A_\mu \mathfrak{a}^\mu = A^i \mathfrak{a}_i, \tag{1}$$

then the covariant derivatives of the covariant and the contravariant components of $\mathbf{A}_{(1)}$ are found to be as follows:

$$A_{\mu\rho} = \frac{\partial A_\mu}{\partial u^\rho} - \{\mu\rho, \epsilon\} A_\epsilon, \tag{2}$$

$$A^i{}_\rho = \frac{\partial A^i}{\partial u^\rho} + \{\epsilon\rho, j\} A^\epsilon. \tag{3}$$

Again, if we have given a tensor $\mathbf{B}_{(2)}$ of the second rank expressed in the forms:

$$\mathbf{B}_{(2)} \qquad B_{\mu\rho}\mathfrak{a}^\rho\mathfrak{a}^\mu = B^{jk}\mathfrak{a}_j\mathfrak{a}_k = B^{j}_{\cdot\rho}\mathfrak{a}_j\mathfrak{a}^\rho, \tag{4}$$

then the covariant derivatives of its covariant, contravariant, and mixed components are found to be as follows:

$$B_{\mu\rho\sigma} = \frac{\partial B_{\mu\rho}}{\partial u^\sigma} - \{\mu\sigma, \epsilon\} B_{\epsilon\rho} - \{\rho\sigma, \epsilon\} B_{\mu\epsilon}, \tag{5}$$

$$B^{jk}_{\cdot\cdot\sigma} = \frac{\partial B^{jk}}{\partial u^\sigma} + \{\epsilon\sigma, j\} B^{\epsilon k} + \{\epsilon\sigma, k\} B^{j\epsilon}, \tag{6}$$

$$B^{j}_{\cdot\rho\sigma} = \frac{\partial B^{j}_{\cdot\rho}}{\partial u^\sigma} + \{\epsilon\sigma, j\} B^{\epsilon}_{\cdot\rho} - \{\rho\sigma, \epsilon\} B^{j}_{\cdot\epsilon}. \tag{7}$$

In successive covariant differentiation of tensor components the order of differentiation is not permutable. This will be shown in detail for the case of the successive covariant differentiation of the covariant components of the vector $\mathbf{A}_{(1)}$ (tensor of the first rank).

Suppose that the tensor of the second rank $\mathbf{B}_{(2)}$ is such that:

$$\mathbf{B}_{(2)} = \nabla\mathbf{A}_{(1)} \qquad \mathfrak{a}^\sigma \frac{\partial}{\partial u^\sigma}(A_\mu \mathfrak{a}^\mu);$$

then $B_{\mu\rho} = A_{\mu\rho}$, and hence by equation (2):

$$B_{\mu\rho} = \frac{\partial A_\mu}{\partial u^\rho} - \{\mu\rho, \epsilon\} A_\epsilon.$$

Introducing this expression for $B_{\mu\rho}$ and the corresponding expressions obtained from it for $B_{\epsilon\rho}$ and $B_{\mu\epsilon}$ in equation (5), making use of the dummy index rule, and noting that $B_{\mu\rho\sigma} = A_{\mu\rho\sigma}$, we obtain:

$$(8)\quad A_{\mu\rho\sigma} = \frac{\partial^2 A_\mu}{\partial u^\sigma \partial u^\rho} - \{\mu\sigma, \epsilon\}\frac{\partial A_\epsilon}{\partial u^\rho} - \{\rho\sigma, \epsilon\}\frac{\partial A_\mu}{\partial u^\epsilon} - \{\mu\rho, \epsilon\}\frac{\partial A_\epsilon}{\partial u^\sigma}$$
$$+\left[\{\rho\sigma, \epsilon\}\{\mu\epsilon, \alpha\} + \{\mu\sigma, \epsilon\}\{\epsilon\rho, \alpha\} - \frac{\partial}{\partial \mu^\sigma}\{\mu\rho, \alpha\}\right]A_\alpha.$$

The expression on the right is called the second covariant derivative of the covariant component A_μ of the vector $\mathbf{A}_{(1)}$, taken first with respect to u^ρ and then with respect to u^σ. It represents the triply covariant typical component of a tensor:

$$(9)\qquad \mathbf{A}_{(3)} = A_{\mu\rho\sigma}\mathfrak{a}^\sigma\mathfrak{a}^\rho\mathfrak{a}^\mu.$$

If the order of covariant differentiation of A_μ with respect to u^ρ and u^σ were permutable, the indices ρ and σ could be interchanged in the expression on the right of equation (8) without alteration of its value. Inspection shows, however, that only the first five terms in general will remain unchanged by this interchange. The order of covariant differentiation is not, therefore, in general permutable.

If $A_{\mu\sigma\rho}$ denote the expression obtained from that on the right of equation (8) by interchange of the indices ρ and σ, we can derive a new tensor, which will be denoted by ${}_{(3)}\mathbf{A}$, by writing:

$$(10)\qquad {}_{(3)}\mathbf{A} = A_{\mu\sigma\rho}\mathfrak{a}^\sigma\mathfrak{a}^\rho\mathfrak{a}^\mu.$$

By subtraction of ${}_{(3)}\mathbf{A}$ from $\mathbf{A}_{(3)}$ yet another tensor of the third rank is obtained, viz:

$$(11)\qquad \mathbf{A}_{(3)} - {}_{(3)}\mathbf{A} = (A_{\mu\rho\sigma} - A_{\mu\sigma\rho})\,\mathfrak{a}^\sigma\mathfrak{a}^\rho\mathfrak{a}^\mu,$$

whose triply covariant components are given by the equation:

$$(12)\quad A_{\mu\rho\sigma} - A_{\mu\sigma\rho} = \left[\{\mu\sigma, \epsilon\}\{\epsilon\rho, \alpha\} - \frac{\partial}{\partial \mu^\sigma}\{\mu\rho, \alpha\}\right.$$
$$\left. - \{\mu\rho, \epsilon\}\{\epsilon\sigma, \alpha\} + \frac{\partial}{\partial u^\rho}\{\mu\sigma, \alpha\}\right]A_\alpha.$$

The covariant derivatives of the components of the fundamental tensor:

$$\mathbf{I} = g_{\alpha\beta}\mathfrak{a}^\alpha\mathfrak{a}^\beta,$$

all vanish identically. For, by equation (4) we have:

$$g_{\alpha\beta\sigma} = \frac{\partial g_{\alpha\beta}}{\partial u^\sigma} - \{\alpha\sigma, \epsilon\}\, g_{\epsilon\beta} - \{\beta\sigma, \alpha\}\, g_{\alpha\epsilon}$$
$$= \frac{\partial g_{\alpha\beta}}{\partial u^\sigma} - [\alpha\sigma, \beta] - [\beta\sigma, \alpha]$$
$$= 0,$$

in virtue of equation (5), Art. 98. In covariant differentiation, therefore, the metrical g-coefficients can be considered as constants.

In virtue of the relationship of the $g^{\alpha\beta}$'s to the $g_{\alpha\beta}$'s implied by equation (5), Art. 94, it follows that the covariant derivatives of the $g^{\alpha\beta}$'s must also vanish identically.

Covariant derivatives of the components of tensors in general reduce to ordinary derivatives when the components of the fundamental tensor are constants, since in this case the Christoffel symbols have zero values. It is largely due to this fact that covariant differentiation is of so much importance in the absolute differential calculus and in the theory of relativity.

If the Christoffel symbols do not vanish on the co-ordinate system which happens to be in use, it is always possible to find another co-ordinate system on which they do vanish and, therefore, for which the distinction between covariant and ordinary differentiation of tensor components disappears, provided the space in which we are operating is Euclidean. If, however, we are dealing with non-Euclidean space, the fact that the Christoffel symbols cannot be made to vanish by change of co-ordinate system and, therefore, that a distinction between covariant and ordinary differentiation of tensor components is inevitable, must be ascribed to a metrical peculiarity (curvature) of the space with which we are dealing.

In addition to the process of covariant differentiation of tensor components there is a corresponding process of contravariant differentiation,[1] but the latter is of comparatively little importance in practical applications of tensor analysis and will not be discussed here.

§116

The Riemann-Christoffel Tensor

The typical component $A_{\mu\rho\sigma} - A_{\mu\sigma\rho}$ of the tensor $\mathbf{A}_{(3)} - {}_{(3)}\mathbf{A}$, expressed by equation (12), Art. 115, must, of course, be covariant with respect to the indices μ, ρ, and σ. It follows that the cofactor of A_α in this equation must be contravariant with respect to the index α and covariant with respect to the indices μ, ρ, and σ. This cofactor, which we shall denote by $R^{\alpha\cdot\cdot\cdot}_{\cdot\mu\rho\sigma}$, must therefore be the typical component of a tensor of the fourth rank:

(1) $$\mathbf{R} = R^{\alpha\cdot\cdot\cdot}_{\cdot\mu\rho\sigma}\mathfrak{a}_\alpha\mathfrak{a}^\mu\mathfrak{a}^\rho\mathfrak{a}^\sigma,$$

[1] Cf. Levi-Civita, The Absolute Differential Calculus, Eng. Tr., p. 149.

where:

(2) $$R^{\alpha\cdot\cdot\cdot}_{\cdot\mu\rho\sigma} = \{\mu\sigma, \epsilon\}\{\epsilon\rho, \alpha\} - \frac{\partial}{\partial u^\sigma}\{\mu\rho, \alpha\} - \{\mu\rho, \epsilon\}\{\epsilon\sigma, \alpha\} + \frac{\partial}{\partial u^\rho}\{\mu\sigma, \alpha\}.$$

This tensor is called the Riemann-Christoffel Tensor. In the case of a V_n for which the Riemann-Christoffel tensor vanishes the order of covariant differentiation is permutable.

Inspection of the typical component of this tensor, given by equation (2), shows that it depends only upon the first and second derivatives of the g-coefficients of the fundamental differential quadratic form; in fact, that it is quadratic in the first and linear in the second derivatives. It therefore ranks with the tensor **I**, defined in Art. 113, as a fundamental tensor of a V_n. Both of these tensors are of fundamental importance in Einstein's Theory of Relativity.

With the aid of equation (b), Art. 106, we can express **R** in the purely reciprocal form:

(3) $$\mathbf{R} = R_{\lambda\mu\rho\sigma}\mathfrak{a}^\lambda\mathfrak{a}^\mu\mathfrak{a}^\rho\mathfrak{a}^\sigma,$$

where:

(4) $$R_{\lambda\mu\rho\sigma} = g_{\lambda\alpha}R^{\alpha\cdot\cdot\cdot}_{\cdot\mu\rho\sigma}.$$

By an easy transformation, the details of which will be omitted, this equation, with the aid of equation (2) above, and of equation (5), Art. 98, and of equation (7), Art. 97, can be reduced to the following form:

(5) $$R_{\lambda\mu\rho\sigma} = -\{\mu\sigma, \epsilon\}[\lambda\rho, \epsilon] + \{\mu\rho, \epsilon\}[\lambda\sigma, \epsilon] + \frac{\partial}{\partial u^\rho}[\mu\sigma, \lambda] - \frac{\partial}{\partial u^\sigma}[\mu\rho, \lambda].$$

Examination of this form for the typical covariant component of the tensor **R** shows that the tensor itself is anti-symmetric with respect to the indices λ and μ, and also with respect to the indices ρ and σ; and, furthermore, that it is symmetric with respect to the double interchange of λ, σ and μ, ρ. Hence:

(6) $$R_{\lambda\mu\rho\sigma} = -R_{\mu\lambda\rho\sigma};$$

(7) $$R_{\lambda\mu\rho\sigma} = -R_{\lambda\mu\sigma\rho};$$

(8) $$R_{\lambda\mu\rho\sigma} = R_{\sigma\rho\mu\lambda}.$$

The following cyclic property of the components of **R** follows directly from equation (5):

(9) $$R_{\lambda\mu\rho\sigma} + R_{\lambda\rho\sigma\mu} + R_{\lambda\sigma\mu\rho} = 0.$$

The general tensor of the fourth rank has n^4 components, but owing to the relationships noted in the preceding paragraph the number (N) of distinct components in the case of the Riemann-Christoffel tensor is far less; in fact, it can be shown[1] that:

$$(10) \qquad N \quad \frac{n^2(n^2 - 1)}{12}.$$

For $n = 4$ the Riemann-Christoffel tensor becomes an important tensor in Einstein's General Theory of Relativity. In this case $N = 20$; there are 21 non-vanishing components with the following indices:[2]

1212	1223	1313	1324	1423	2323	2424
1213	1224	1314	1334	1424	2324	2434
1214	1234	1323	1414	1434	2334	3434;

these components are distinct except for the relation:

$$R_{1234} - R_{1324} + R_{1423} = 0.$$

When $n = 3$ there are 6 distinct components, obtainable from the above table by excluding those numbers which contain the figure 4.

When $n = 2$ (the case of a surface) there is but one distinct component, which can be taken as R_{1212}.

When it is possible to find co-ordinates for a V_n such that the fundamental differential quadratic form will reduce to a sum of squares, the space must be Euclidean. Now this will be possible if the g-coefficients of the form are all constants, as can be shown[3] by well known algebraic methods. But in this case the Riemann-Christoffel tensor must vanish throughout the space as can be seen by inspection of equation (5), noting that the Christoffel symbols must vanish. The vanishing of the Riemann-Christoffel tensor throughout a V_n is therefore a necessary condition that the space shall be a Euclidean Manifold; it can be proved that it is also a sufficient condition.[4]

Since there is no curvature of Euclidean space, it is therefore evident that the measure of curvature of a general V_n must involve in some way the Riemann-Christoffel tensor.

[1] Cf. Levi-Civita, The Absolute Differential Calculus, Eng. tr., p. 181.
[2] Cf. Eddington, The Mathematical Theory of Relativity, p. 72.
[3] Cf. Eddington, The Mathematical Theory of Relativity, p. 13.
[4] Cf. Eddington, The Mathematical Theory of Relativity, p. 72.

§117

The Ricci-Einstein Tensor

The Riemann-Christoffel tensor:

$$\mathbf{R} = R_{\lambda\mu\rho\sigma}\mathfrak{a}^\lambda\mathfrak{a}^\mu\mathfrak{a}^\rho\mathfrak{a}^\sigma,$$

when contracted (see Art. 112) with respect to the indices μ and ρ, gives a tensor of the second rank:

$$\begin{aligned}\mathbf{G} &= R_{\lambda\mu\rho\sigma}\mathfrak{a}^\mu \cdot \mathfrak{a}^\rho\mathfrak{a}^\lambda\mathfrak{a}^\sigma\\ &= R_{\lambda\mu\rho\sigma}g^{\mu\rho}\mathfrak{a}^\lambda\mathfrak{a}^\sigma.\end{aligned}$$

It can be written in the following form:

(1) $$\mathbf{G} = G_{\lambda\sigma}\mathfrak{a}^\lambda\mathfrak{a}^\sigma,$$

where:

(2) $$G_{\lambda\sigma} = g^{\mu\rho}R_{\lambda\mu\rho\sigma}.$$

Since the interchange of the dummy indices μ and ρ does not alter the component $G_{\lambda\sigma}$ it follows from equation (8), Art. 116, that the tensor $\mathbf{G}$ is symmetric.

This tensor was first noticed by Ricci in connection with his study of the local curvatures of a metric manifold but, on account of its importance in relativity theory, it is now commonly known as the Ricci-Einstein Tensor.

Upon contraction of this tensor, we obtain a tensor of zero rank, that is, a scalar invariant:

(3) $$\begin{aligned}K_0 &= G_{\lambda\sigma}\mathfrak{a}^\lambda \cdot \mathfrak{a}^\sigma,\\ &= g^{\lambda\sigma}G_{\lambda\sigma}.\end{aligned}$$

This invariant is of importance in connection with the measure of the curvature of a V_n, as will be shown in Art. 119 for the particular case of a V_2.

For the particular case of a V_2 (a surface), as was seen in Art. 116, the Riemann-Christoffel tensor has but one distinct component. If we take this to be R_{1212}, the tensor itself, with the aid of the relations (6), (7), and (8), Art. 116, can be written in the form:

(4) $$\mathbf{R} = R_{1212}(\mathfrak{a}^1\mathfrak{a}^2 - \mathfrak{a}^2\mathfrak{a}^1)(\mathfrak{a}^1\mathfrak{a}^2 - \mathfrak{a}^2\mathfrak{a}^1).$$

Upon contraction of this tensor we get the Ricci-Einstein tensor for the particular case of a V_2 (a surface) in the forms:

$$\begin{aligned}\mathbf{G} &= R_{1212}(\mathfrak{a}^1\mathfrak{a}^2 - \mathfrak{a}^2\mathfrak{a}^1)\cdot(\mathfrak{a}^1\mathfrak{a}^2 - \mathfrak{a}^2\mathfrak{a}^1)\\ &= R_{1212}[-g^{22}\mathfrak{a}^1\mathfrak{a}^1 + g^{12}(\mathfrak{a}^1\mathfrak{a}^2 + \mathfrak{a}^2\mathfrak{a}^1) - g^{11}\mathfrak{a}^2\mathfrak{a}^2]\\ &- \frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{12}}[g_{11}\mathfrak{a}^1\mathfrak{a}^1 + g_{12}(\mathfrak{a}^1\mathfrak{a}^2 + \mathfrak{a}^2\mathfrak{a}^1) + g_{22}\mathfrak{a}^2\mathfrak{a}^2],\end{aligned}$$

with the aid of the first of equations (9), Art. 94. Consequently, we can write:

$$G = G_{\lambda\mu}\mathfrak{a}^\lambda\mathfrak{a}^\mu, \qquad (\lambda, \mu = 1, 2), \tag{5}$$

where:

$$G_{\lambda\mu} = -\frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{12}}\, g_{\lambda\mu}, \qquad (\lambda, \mu = 1, 2). \tag{6}$$

Upon contraction of the tensor in the form last given we get the scalar invariant:

$$\begin{aligned} K_0 &= G_{\lambda\mu}\mathfrak{a}^\lambda \cdot \mathfrak{a}^\mu \\ &= G_{\lambda\mu}g^{\lambda\mu}, \qquad (\lambda, \mu = 1, 2), \end{aligned}$$

or, upon taking account of equation (6):

$$K_0 = -\frac{2\, R_{1212}}{g_{11}g_{22} - g_{12}g_{12}}. \tag{7}$$

In Art. 119 it will be shown that the negative of this invariant furnishes a measure of the Gaussian curvature at any point of an ordinary surface.

§118

The Change of a Vector in a V_n Produced by its Parallel Displacement around an Infinitesimal Quadrilateral

In Euclidean space there is no change produced in a vector by its parallel displacement around any closed cycle, since it thereby experiences no change in magnitude or direction. But for a corresponding displacement of a vector in a general V_n this will not be the case, for, as we know, although its magnitude would remain constant its direction would in general be changed. We now propose to find an expression for the change in a V_n vector produced by its parallel displacement once around an infinitesimal quadrilateral.[1]

Consider an infinitesimal quadrilateral $PQRS$ in a V_n, and suppose a co-ordinate system so chosen as to make the sides of the quadrilateral elements of the co-ordinate lines of two of the co-ordinates, say u^ρ and u^σ. At the vertices of the quadrilateral these co-ordinates will have values which are indicated as follows:

$$P(u^\rho, u^\sigma), \quad Q(u^\rho, u^\rho + du^\rho), \quad R(u^\rho + du^\rho, u^\sigma + \delta u^\sigma), \quad S(u^\rho, u^\sigma + \delta u^\sigma).$$

It is first desired to investigate the changes in the field of a vector point function **A** met with in passing around the infinitesimal

[1] Cf. Eddington, The Mathematical Theory of Relativity, p. 69.

quadrilateral. To do this, we shall first find the increase in the typical covariant component A_μ of the vector **A** in passing in the direction $P \rightarrow Q \rightarrow R \rightarrow S$ around the quadrilateral.

In passing from P to Q the vector must experience an absolute infinitesimal change which is given by equation (13), Art. 100, from which we infer that the corresponding increase in its typical covariant component A_μ is given by the expression:

$$\left[\frac{\partial A_\mu}{\partial u^\rho} - \{\mu\rho, \epsilon\}\, A_\epsilon\right] du^\rho, \qquad \text{(not summed as regards } \rho\text{)},$$

which is the covariant derivative $(A_{\mu\rho})$ of A_μ, evaluated at P. Therefore, the total increase in A_μ, say ΔA_μ, will be given by the following expression:

$$\Delta A_\mu = (A_{\mu\rho})_P\, du^\rho + (A_{\mu\sigma})_Q\, \delta u^\sigma - (A_{\mu\sigma})_P\, \delta u^\sigma - (A_{\mu\rho})_S\, du^\rho,$$

where no summation is implied, and where the quantities in parentheses denote covariant derivatives to be evaluated at the points indicated by corresponding subscripts. Now:

$$(A_{\mu\sigma})_Q = (A_{\mu\sigma})_P + \text{covariant derivative of } (A_{\mu\sigma})_P \times du^\rho,$$
$$(A_{\mu\rho})_S = (A_{\mu\rho})_P + \text{covariant derivative of } (A_{\mu\rho})_P \times \delta u^\sigma.$$

Hence:

(1) $$\Delta A_\mu = [A_{\mu\sigma\rho} - A_{\mu\rho\sigma}]\, du^\rho \delta u^\sigma,$$

where the quantities in brackets denote second covariant derivatives of A_μ to be evaluated at P.

The quantity ΔA_μ represents the amount by which A_μ must be increased in order that the cycle shall begin and end with the same vector **A**.

Consider next a **parallel** displacement of the vector **A** in the direction $P \rightarrow Q \rightarrow R \rightarrow S$ around the infinitesimal quadrilateral, beginning at P. In this displacement no **absolute** change in the covariant component A_μ is allowed, and upon completion of the cycle the displaced vector will have a value at P which differs from that of **A**; there will be a difference, say DA_μ, in the value of A_μ for the displaced vector and for the original vector of an amount obviously equal to $-\Delta A_\mu$. Hence:

(2) $$DA_\mu = [A_{\mu\rho\sigma} - A_{\mu\sigma\rho}]\, du^\rho \delta u^\sigma, \qquad \text{(not summed)}.$$

With the aid of this result the increase, say $D\mathbf{A}$, in the vector **A** due to its parallel displacement in the direction $P \rightarrow Q \rightarrow R \rightarrow S$ once around the infinitesimal quadrilateral can be expressed as follows:

(3) $$D\mathbf{A} = DA_\mu \mathbf{a}^\mu = [A_{\mu\rho\sigma} - A_{\mu\sigma\rho}]\, du^\rho \delta u^\sigma \mathbf{a}^\mu, \qquad \text{(summed as regards } \mu\text{)}.$$

This equation has been derived on the assumption of a special co-ordinate system, but, since it is a tensor equation, it must be valid on all co-ordinate systems; furthermore, the result expressed by this equation must be valid for an infinitesimal circuit of any shape, since co-ordinates can always be found such that the circuit may be considered as constituted of elements of co-ordinate lines of two co-ordinates.

We now suppose the special co-ordinate system to be replaced by a general u-co-ordinate system. In passing from the vertex P to the vertex Q, or from the vertex P to the vertex S, of the infinitesimal quadrilateral $PQRS$, all the co-ordinates will vary in general by amounts which in the first case will be designated by the symbol d and in the second case by the symbol δ. The tensor equation (3) will be valid on the new system, provided summation with respect to *all* the indices on the right is understood.

It was pointed out in Art. 116 that:

$$A_{\mu\rho\sigma} - A_\mu \qquad R^{\alpha\cdots}_{\cdot\mu\rho\sigma}A_\alpha,$$

where $R^{\alpha\cdots}_{\cdot\mu\rho\sigma}$ denotes the typical component of the Riemann-Christoffel tensor, and we can therefore write:

$$\begin{aligned} D\mathbf{A} &= R^{\alpha\cdots}_{\cdot\mu\rho\sigma}du^\rho\delta u^\sigma\mathbf{a}^\mu A \\ &= R_{\lambda\mu\rho\sigma}du^\rho\delta u^\sigma\mathbf{a}^\mu A^\lambda \\ &= R_{\lambda\mu\rho\sigma}du^\rho\delta u^\sigma\mathbf{a}^\mu\mathbf{a}^\lambda\cdot\mathbf{A}, \quad \text{(summation understood)}, \end{aligned}$$

or, since the Riemann-Christoffel tensor is anti-symmetric with respect to the indices λ and μ:

$$D\mathbf{A} = -R_{\lambda\mu\rho\sigma}du^\rho\delta u^\sigma\mathbf{a}^\lambda\mathbf{a}^\mu\cdot\mathbf{A}. \tag{4}$$

If we write:

$$\mathbf{\Phi} = R_{\lambda\mu\rho\sigma}du^\rho\delta u^\sigma\mathbf{a}^\lambda\mathbf{a}^\mu, \tag{5}$$

then:

$$D\mathbf{A} = -\mathbf{\Phi}\cdot\mathbf{A}. \tag{6}$$

The quantity $\mathbf{\Phi}$ must be a tensor, since it is both covariant and contravariant with respect to each of the indices λ, μ, ρ, σ. This tensor is of the second rank, and is anti-symmetric with respect to interchange of the indices λ and μ, since, as shown in Art. 116, the factor $R_{\lambda\mu\rho\sigma}$ is anti-symmetric with respect to this interchange. We can, therefore, also write:

$$\mathbf{\Phi} = -R_{\lambda\mu\rho\sigma}du^\rho\delta u^\sigma\mathbf{a}^\mu\mathbf{a}^\lambda. \tag{7}$$

By combination of equations (5) and (7) we get:

$$\mathbf{\Phi} = \frac{1}{2} R_{\lambda\mu\rho\sigma} du^\rho \delta u^\sigma (\mathfrak{a}^\lambda \mathfrak{a}^\mu - \mathfrak{a}^\mu \mathfrak{a}^\lambda). \tag{8}$$

If the signs d and δ in this expression be interchanged, the only effect is to interchange the rôles of the indices ρ and σ, so that:

$$\mathbf{\Phi} = \frac{1}{2} R_{\lambda\mu\sigma\rho} du^\sigma \delta u^\rho (\mathfrak{a}^\lambda \mathfrak{a}^\mu - \mathfrak{a}^\mu \mathfrak{a}^\lambda), \tag{9}$$

or since, as shown in Art. 116, the quantity $R_{\lambda\mu\rho\sigma}$ is anti-symmetric with respect to the indices ρ and σ:

$$\mathbf{\Phi} = -\frac{1}{2} R_{\lambda\mu\rho\sigma} du^\sigma \delta u^\rho (\mathfrak{a}^\lambda \mathfrak{a}^\mu - \mathfrak{a}^\mu \mathfrak{a}^\lambda). \tag{10}$$

By combination of equations (8) and (10) we get:

$$\mathbf{\Phi} = \frac{1}{4} R_{\lambda\mu\rho\sigma} (du^\rho \delta u^\sigma - \delta u^\rho du^\sigma)\,(\mathfrak{a}^\lambda \mathfrak{a}^\mu - \mathfrak{a}^\mu \mathfrak{a}^\lambda). \tag{11}$$

Now, for the infinitesimal position-vectors ($d\mathbf{s}$, $\delta\mathbf{s}$) of the points Q and S with respect to P we have:

$$d\mathbf{s} = \mathfrak{a}_\rho du^\rho, \qquad \delta\mathbf{s} = \mathfrak{a}_\sigma \delta u^\sigma, \tag{12}$$

and hence:

$$\begin{aligned} du^\rho &= \mathfrak{a}^\rho \cdot d\mathbf{s}, & \delta u^\sigma &= \mathfrak{a}^\sigma \cdot \delta\mathbf{s}, \\ du^\sigma &= \mathfrak{a}^\sigma \cdot d\mathbf{s}, & \delta u^\rho &= \mathfrak{a}^\rho \cdot \delta\mathbf{s}, \end{aligned}$$

so that:

$$du^\rho \delta u^\sigma - \delta u^\rho du^\sigma = \mathfrak{a}^\rho \cdot (d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s} d\mathbf{s}) \cdot \mathfrak{a}^\sigma. \tag{13}$$

By combination of equations (11) and (13) we get:

$$\mathbf{\Phi} = \frac{1}{4} R_{\lambda\mu\rho\sigma} \mathfrak{a}^\rho \cdot (d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s} d\mathbf{s}) \cdot \mathfrak{a}^\sigma (\mathfrak{a}^\lambda \mathfrak{a}^\mu - \mathfrak{a}^\mu \mathfrak{a}^\lambda). \tag{14}$$

Using this expression in equation (6), we find finally that:

$$D\mathbf{A} = -\left[\frac{1}{4} R_{\lambda\mu\rho\sigma} \mathfrak{a}^\rho \cdot (d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s} d\mathbf{s}) \cdot \mathfrak{a}^\sigma (\mathfrak{a}^\lambda \mathfrak{a}^\mu - \mathfrak{a}^\mu \mathfrak{a}^\lambda)\right] \cdot \mathbf{A}. \tag{15}$$

In the particular case of a surface this equation reduces to the following one:

$$D\mathbf{A} = -[R_{1212} \mathfrak{a}^1 \cdot (d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s} d\mathbf{s}) \cdot \mathfrak{a}^2 (\mathfrak{a}^1 \mathfrak{a}^2 - \mathfrak{a}^2 \mathfrak{a}^1)] \cdot \mathbf{A}. \tag{16}$$

In this case, from equations (12), with the aid of the second of equations (2), Art. 94, we have:

$$\begin{aligned} d\mathbf{s} &= \mathfrak{a}_\rho du^\rho = g_{\rho\alpha} \mathfrak{a}^\alpha du^\rho, & (\rho, \alpha &= 1, 2), \\ \delta\mathbf{s} &= \mathfrak{a}_\sigma \delta u^\sigma = g_{\sigma\beta} \mathfrak{a}^\beta \delta u^\sigma, & (\sigma, \beta &= 1, 2), \end{aligned}$$

and from these equations it follows directly that the dyadic $d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s}d\mathbf{s}$ differs from the dyadic $\mathfrak{a}^1\mathfrak{a}^2 - \mathfrak{a}^2\mathfrak{a}^1$ by a scalar factor only. Hence, in equation (16) it is permissible to interchange the positions of these dyadics, so that:

$$\begin{aligned} D\mathbf{A} &= -[R_{1212}\mathfrak{a}^1 \cdot (\mathfrak{a}^1\mathfrak{a}^2 - \mathfrak{a}^2\mathfrak{a}^1) \cdot \mathfrak{a}^2(d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s}d\mathbf{s})] \cdot \mathbf{A} \\ &= \quad [R_{1212}(\mathfrak{a}^1 \cdot \mathfrak{a}^1\mathfrak{a}^2 \cdot \mathfrak{a}^2 - \mathfrak{a}^1 \cdot \mathfrak{a}^2\mathfrak{a}^1 \cdot \mathfrak{a}^2)\,(d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s}d\mathbf{s})] \cdot \mathbf{A} \\ &= \quad [R_{1212}(g^{11}g^{22} - g^{12}g^{12})\,(d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s}d\mathbf{s})] \cdot \mathbf{A}, \end{aligned}$$

or, making use of the first of equations (9), Art. 94:

$$D\mathbf{A} = -\frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{12}} \qquad - \delta\mathbf{s}d\mathbf{s}) \cdot \mathbf{A}. \tag{17}$$

In this expression for the change produced in a surface vector $\mathbf{A}$ by its parallel displacement once around the infinitesimal quadrilateral having two adjacent sides determined by the infinitesimal surface vectors $d\mathbf{s}$ and $\delta\mathbf{s}$, the scalar factor

$$\frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{12}}$$

is by equation (7), Art. 117, equal to the negative of the scalar invariant K_0, obtained by the contraction of the Einstein tensor for a surface, and depends solely upon the metrical properties of the surface; the dyadic

$$- \delta\mathbf{s}d\mathbf{s}$$

depends solely upon the geometrical configuration of the initial point P and the terminal points Q, S of $d\mathbf{s}$ ond $\delta\mathbf{s}$.

§119

Calculation of the Gaussian Curvature of a Surface

The result expressed by equation (17) of the preceding article enables us, as will be shown, to calculate the Gaussian curvature at a point of an ordinary surface.

It was shown in Art. 102 that the Gaussian curvature of a surface at a point P can be expressed as follows:

$$K = \frac{D\alpha}{D\sigma}, \tag{1}$$

where $D\alpha$ represents the angular change in direction of a surface vector $\mathbf{A}$ which would be produced in its parallel displacement once around the contour bounding an infinitesimal surface element of area $D\sigma$, the contour being supposed to pass through the point P,

and $D\alpha$ being supposed reckoned positive in the rotary sense of the displacement of the vector around the contour. If $D\mathbf{A}$ denote the change in the vector $\mathbf{A}$ due to such a displacement, then $D\alpha$ must be the angle between the vectors $\mathbf{A}$ and $\mathbf{A} + D\mathbf{A}$ at P.

We now suppose the element of area of the surface to be that of an infinitesimal parallelogram whose sides are determined by the infinitesimal surface vectors:

$$d\mathbf{s} = \mathbf{a}_1 du^1, \qquad \delta\mathbf{s} = \mathbf{a}_2 \delta u^2,$$

with vertices at the points

$$P(u^1, u^2), \quad Q(u^1 + du^1, u^2), \quad R(u^1 + du^1, u^2 + \delta u^2), \quad S(u^1, u^2 + \delta u^2).$$

Since $\mathbf{A}$ would undergo no change in magnitude in a parallel displacement around this parallelogram, therefore $D\mathbf{A}$ must be a surface-vector perpendicular to $\mathbf{A}$.

If $\mathbf{n}$ denote a unit vector in the direction and sense of the vector product $d\mathbf{s} \times \delta\mathbf{s}$, then $\mathbf{n} \times \mathbf{A}$ will be a surface-vector at P which is perpendicular to $\mathbf{A}$ (and therefore parallel to $D\mathbf{A}$) in the direction of α increasing. Hence we shall have:

(2) $$D\mathbf{A} = D\alpha\mathbf{n} \times \mathbf{A}.$$

Furthermore, we can write:

$$d\mathbf{s} \times \delta\mathbf{s} = \mathbf{n} D\sigma.$$

By vector multiplication of this equation by $\mathbf{A}$ we get:

$$(d\mathbf{s} \times \delta\mathbf{s}) \times \mathbf{A} = \mathbf{n} \times \mathbf{A} D\sigma,$$

and hence:

(3) $$(d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s} d\mathbf{s}) \cdot \mathbf{A} = -\mathbf{n} \times \mathbf{A} D\sigma.$$

From equations (2) and (3) it follows that:

(4) $$D\mathbf{A} = -\frac{D\alpha}{d\sigma}(d\mathbf{s}\delta\mathbf{s} - \delta\mathbf{s} d\mathbf{s}) \cdot \mathbf{A}.$$

Comparison of this equation for $D\mathbf{A}$ with that given by equation (17), Art. 118, gives:

(5) $$\frac{D\alpha}{D\sigma} = \frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{12}}$$

Finally, comparison of this equation with equation (1) gives for the Gaussian curvature of the surface at the point P:

(6) $$K = \frac{R_{1212}}{g_{11}g_{22} - g_{12}g_{12}}.$$

This expression for the Gaussian curvature of the surface is equal to the negative of the invariant K_0 obtained by contraction of the Ricci-Einstein tensor of the surface in Art. 117.

EXERCISES ON CHAPTER X

1. The transformation equations for a set of n^2 scalar quantities are known to be of the same forms as for the components of a V_n-tensor of the second rank; show that the quantities must be the components of a V_n tensor of the second rank.

2. If **T** and **S** are two V_n-tensors of the second rank, show that $T_{ij}S^{ij}$ (summation understood) is an invariant.

3. If **S** is an arbitrary tensor of the second rank and if $T_{ij}S^{ij}$ (summation understood) is an invariant, show that the T_{ij}'s are the components of a tensor.

4. It is known that the product

$$A(i, \lambda)\, B^i, \qquad \text{(summation understood),}$$

where $A(i, \lambda)$ is a scalar expression involving the indices i and λ and B^i is the component of an arbitrary vector, is the component of a tensor; show that the quantity $A(i, \lambda)$ must itself be a covariant component of a tensor.

5. Show that the transformation equations for the components of tensors satisfy the fundamental group property.

6. If a, b, c denote the numbers of distinct components of a V_n-tensor of the third rank for which, respectively, the indices are all different, two only are different, all are the same, show that for

a symmetric tensor	an anti-symmetric tensor
$a = \dfrac{n(n-1)(n-2)}{6}$,	$a = \dfrac{n(n-1)(n-2)}{6}$,
$b = n(n-1)$,	$b = 0$,
$c = n$,	$c = 0$.

7. If **A** and **B** are two vectors associated with a point of a V_n, show that $A_\lambda B_\mu - A_\mu B_\lambda$ are components of an anti-symmetric tensor.

8. Show that the differential invariants ∇U and $\nabla U \cdot \nabla U$, where U is a scalar point function in a V_3, can be expressed in the forms:

$$\nabla U \quad 1 \begin{vmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & 0 \\ g_{11} & g_{12} & g_{13} & \dfrac{\partial U}{\partial u^1} \\ g_{21} & g_{22} & g_{23} & \dfrac{\partial U}{\partial u^2} \\ g_{31} & g_{32} & g_{33} & \dfrac{\partial U}{\partial u^3} \end{vmatrix}, \qquad \nabla U \cdot \nabla U = \frac{1}{g} \begin{vmatrix} \dfrac{\partial U}{\partial u^1} & \dfrac{\partial U}{\partial u^2} & \dfrac{\partial U}{\partial u^3} & 0 \\ g_{11} & g_{12} & g_{13} & \dfrac{\partial U}{\partial u^1} \\ g_{21} & g_{22} & g_{23} & \dfrac{\partial U}{\partial u^2} \\ g_{31} & g_{32} & g_{33} & \dfrac{\partial U}{\partial u^3} \end{vmatrix}$$

9. If $A^{i\cdot}_{\cdot\mu}$ is a mixed component of a symmetric tensor, show that

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial u^i}(\sqrt{g}A^{i\cdot}_{\cdot\mu}) + \frac{1}{2}\frac{\partial g^{\alpha\beta}}{\partial u^\mu}A_{\alpha\beta}, \quad \text{(summation understood)},$$

is the component of a vector.

10. If A^{ij} is a contravariant component of an anti-symmetric tensor, show that

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial u^j}(\sqrt{g}A^{ij}), \quad \text{(summation understood)},$$

is the component of vector.

11. Write down a set of rules which will serve as an adequate guide in covariant differentiation. (See Eddington, Math. Theory of Relativity, p. 65.)

12. Show that the Gaussian curvature K of a surface can be expressed as follows:

$$K = \frac{1}{2\sqrt{g}}\left\{\frac{\partial}{\partial u^1}\left[\frac{g_{12}}{g_{11}\sqrt{g}}\frac{\partial g_{11}}{\partial u^2} - \frac{1}{\sqrt{g}}\frac{\partial g_{22}}{\partial u^1}\right]\right.$$
$$\left. - \frac{\partial}{\partial u^2}\left[\frac{2}{\sqrt{g}}\frac{\partial g_{12}}{\partial u^1} - \frac{1}{\sqrt{g}}\frac{\partial g_{11}}{\partial u^2} - \frac{g_{12}}{g_{11}\sqrt{g}}\frac{\partial g_{11}}{\partial u^1}\right]\right\}$$

REFERENCES

The list of references given below by no means constitutes a complete bibliography of the subject matter of the present book, but it is probably sufficiently comprehensive for the needs of its readers. Specific acknowledgments of the author's indebtedness to all of the books of this list of which he has made use is not possible, but those of Gibbs-Wilson, Levi-Civita, and Lagally demand special mention:

Abraham und Föppl. Theorie der Elektrizität, Bd. I, 1904.

Ames and Murnaghan. Theoretical Mechanics, 1929.

Blaschke, W. Vorlesungen über Differentialgeometrie, 1924.

Coffin, J. G. Vector Analysis, 1911.

Eddington, A. S. The Mathematical Theory of Relativity, 1923.

Eisenhart, L. P. Riemannian Geometry, 1926.

Gans, R. Einführung in die Vektoranalysis, 1909.

Gibbs-Wilson. Vector Analysis, 1901.

Grassmann, Hermann. Die Ausdehnungslehre von 1862.

Hamilton, W. R. Elements of Quaternions, 1866.

Heaviside, Oliver. Electromagnetic Theory, Vol. I, p. 132, 1893.

Hessenberg, Gerhard. Vektorielle Begründung der Differentialgeometrie, Math. Ann. 78, s. 187, 1916.

Ignatowsky, W. Die Vektoranalysis, 1909.

Lagally, M. Vektor-Rechnung, 1928.

Levi-Civita, T. The Absolute Differential Calculus, Eng. Tr., 1927.

Murnaghan, F. D. Vector Analysis and the Theory of Relativity, 1922.

Radakovic und Lense. Vektor- und Tensorrechnung, Handbuch der Physik, Bd. III, s. 182, 1928.

Runge, C. Vector Analysis, Eng. Tr., 1919.

Schouten, J. A. Der Ricci-Kalkül, 1924.

Silberstein, L. Elements of Vector Algebra, 1919.

Speilrein, I. Vektorrechnung, 1926.

Tait, P. G. An Elementary Treatise on Quaternions, 1867.

Weatherburn, C. E. Elementary and Advanced Vector Analysis, 1926.

INDEX

THE NUMBERS REFER TO PAGES